Environmental Research 225 (2023) 115598

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Multi-pollutant case-crossover models of all-cause and cause-specific mortality and hospital admissions by age group in 47 Canadian cities

Guowen Huang [a,b], Patrick Brown [a,b], Hwashin Hyun Shin [c,d,*]

[a] Department of Statistical Sciences, University of Toronto, Toronto, ON, Canada
[b] Centre for Global Health Research, St Michael's Hospital, Toronto, ON, Canada
[c] Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, Canada
[d] Department of Mathematics and Statistics, Queen's University, Kingston, ON, Canada

ARTICLE INFO

Handling Editor: Jose L Domingo

Keywords:
Air pollution
Case-crossover model
Hospitalization
Mortality
Short-term exposure

ABSTRACT

Most of the existing epidemiological studies have investigated adverse health effects of multiple air pollutants for a limited number of cities, thus the evidence of the health impacts is limited and it is challenging to compare these results because of different modeling approaches and potential publication bias. In this paper, we expand the number of Canadian cities, with the use of the most recent available health data. A multi-pollutant model in a case-crossover design is used to investigate the short-term impacts of air pollution on various health outcomes in 47 Canadian main cities, comparing three age groups (all-age, senior (age 66+), non-senior). The main findings are that a 14 ppb increase of O_3 was associated with a 0.17%–2.78% (0.62%–1.46%) increase in the odds of all-age respiratory mortality (hospitalization). A 12.8 ppb increase of NO_2 was associated with a 0.57%–1.47% (0.68%–1.86%) increase in the odds of all-age (non-senior) respiratory hospitalization. A 7.6 $\mu g m^{-3}$ increase of $PM_{2.5}$ was associated with a 0.19%–0.69% (0.33%–1.1%) increase in the odds of all-age (non-senior) respiratory hospitalization.

1. Introduction

Air pollution remains an important worldwide public health concern. The World Health Organization (WHO, 2021) reports that 99% of the world's population was living in places where the WHO air quality guidelines levels were not met in 2019, and that ambient (outdoor) air pollution in both cities and rural areas was estimated to cause 4.2 million premature deaths worldwide in 2016. Tracking backwards, the health impacts of air pollution exposure has been widely recognised since the 1950's, as a result of the London smog in December 1952, which was estimated to have resulted in more than 3000 excess deaths compared with previous years (Bell and Davis, 2001). Recently, Di et al. (2017) studied the association of short-term exposure to air pollution with mortality in older adults and found that fine particulate matter ($PM_{2.5}$) and ground-level ozone (O_3) were linked to higher risk of premature death among the elderly in the United States. In summary, the adverse human health effects associated with exposure to air pollutants have been well documented. For example, mortality from respiratory and heart diseases was significantly related to the level of air pollution (see e.g., Hodgson, 1970), while all-cause mortality was found to be significantly associated with O_3 levels (see e.g., Anderson et al., 1996; Huang et al., 2022).

According to the existing literature in epidemiological studies, the adverse effects on health associated with air pollution exposure can be considered in two ways, the long-term and short-term effects. When estimating long-term effects, either cohort or spatio-temporal ecological designs can be used (see e.g., Pope et al., 2002; Hoek et al., 2001; Haining et al. 2010; Lawson et al., 2012; Rushworth et al., 2014). For estimating short-term effects, one of the traditional methods is the overdispersed Poisson time series models based on an ecological study design (see e.g., Stieb et al., 2008; Liu et al., 2019; Du et al., 2020), which usually analyze city-level associations and then pool them to represent national associations. Another competing method for such studies is the case-crossover analysis (see e.g., Gasparrini et al., 2015; H. Chen et al., 2016; Huang et al., 2022). In a case-crossover model, the environment on the date of each health outcome (case day) is contrasted with the environment on one or more 'control days', therefore the individual-level covariates (e.g., age, sex, socioeconomic status, pre-existing health conditions) are not confounders as they remain constant on both case and control days.

* Corresponding author. Environmental Health Science and Research Bureau, Health Canada, Ottawa, ON, Canada.
E-mail address: hwashin.shin@hc-sc.gc.ca (H.H. Shin).

https://doi.org/10.1016/j.envres.2023.115598
Received 20 November 2022; Received in revised form 24 February 2023; Accepted 28 February 2023
Available online 1 March 2023

Most of the existing epidemiological studies in Canada have investigated the air pollution health effects for a limited number of cities, either just a specific city/area or a few of them at once. Thus, the evidence of the health impacts is limited and it is challenging to compare these results because of different modeling approaches. For example, Stieb et al. (2008) used a time-series model to analyze the association between air pollution and daily mortality in 12 Canadian cities, based on which the authors proposed a new air quality health index (AQHI). They reported a 1.69% change in daily mortality associated with a 12.8 μgm^{-3} increase of $PM_{2.5}$; a 2.08% change in daily mortality associated with a 33.6 ppb (parts per billion) increase of nitrogen dioxide (NO_2); a 1.82% change in daily mortality associated with a 29.9 ppb increase of O_3. Huang et al. (2022) used a case-crossover design to investigate how the short-term simultaneous exposure to multiple pollutants ($PM_{2.5}$, NO_2 and O_3) related to daily mortality in 4 big Canadian cities. Their results suggested a 3.88% increase in daily mortality and a 0.76% increase in daily morbidity associated with a 10 ppb increase of O_3, and a 2.13% increase in circulatory morbidity associated with 10 ppb increase of NO_2. In Liu et al. (2019), the author analyzed the daily data on mortality and air pollution in 652 cities from 24 countries or regions, among which, 25 Canadian cities were included. The study suggested a 1.7% change in all-cause mortality per 10 μgm^{-3} increase in 2-Day moving average concentrations of $PM_{2.5}$ in Canada. These existing results are not directly comparable due to their different modeling approaches and relative risks related to various increasing units of pollutants. In this paper, we expand the number of Canadian cities, with the use of most recent available health data (mortality from 2001 to 2015 and hospitalization from 2001 to 2018). Specifically, we apply a multi-pollutant model in a case-crossover design to investigate the short-term impacts of air pollution on health in 47 Canadian Census Divisions (CDs), each of which has population larger than 40,000, covering about 60% of total Canadian population. The health outcomes include three types of age groups (all-age, senior (age 65+), non-senior) of three causes (all-cause, respiratory, circulatory) of mortality and hospitalization, which turn out to be a total of 18 health outcomes. The findings in the paper could potentially provide information for making future modifications to the AQHI in Canada.

The remainder of this article is organized as follows. Section 2 introduces datasets and statistical methods used in this study, while Section 3 presents the main results, including the city-specific and national health effects of air pollution. Section 4 is a concluding discussion. More results in detail are reported as Supplementary Materials.

2. Materials and methods

2.1. Data source and summary

The study population included 47 CDs in nine provinces across Canada as shown in Fig. 1, covering about 60% of the target population (total Canadian population). The 47 CDs were selected mainly based on population size (at least 40,000 residents) and data availability of air pollution measured by the ground-monitoring stations. For each CD, daily data of health outcomes, air pollution concentrations, and temperature were collected.

There were two types of health outcome data: mortality from 2001 to 2015 and hospitalization from 2001 to 2018. Both data sets included information on age and cause classified by the International Classification of Diseases 10th Revision (ICD-10): by age group (all-age, non-senior (between 1 and 65, inclusive), and senior (over 65)); and by cause (non-accidental all-cause (A00-R99), circulatory (ICD-10, I00–I99) and respiratory (ICD-10, J00-J99)). Note that the all-cause hospitalization data are available from April 1st, 2001 rather than January 1st, 2001. Daily counts of mortality were obtained from the Canadian Vital Statistics Death Database managed by Statistics Canada, and hospitalization data were obtained from the Canadian Institute for Health Information (source: Hospital Morbidity Database and Discharge Abstract Database). For hospitalization data, we included only acute hospital admissions, which would effectively exclude all other types of admissions such as inpatient rehab, day surgery, etc. Taking age group and cause together, this study provides age-cause-specific health risk estimates.

Three-hour maximum concentrations (3hMax) of $PM_{2.5}$, NO_2 and O_3 were obtained from 2001 to 2018 for each CD. All air pollution data were obtained with a low missing rate from the National Air Pollution

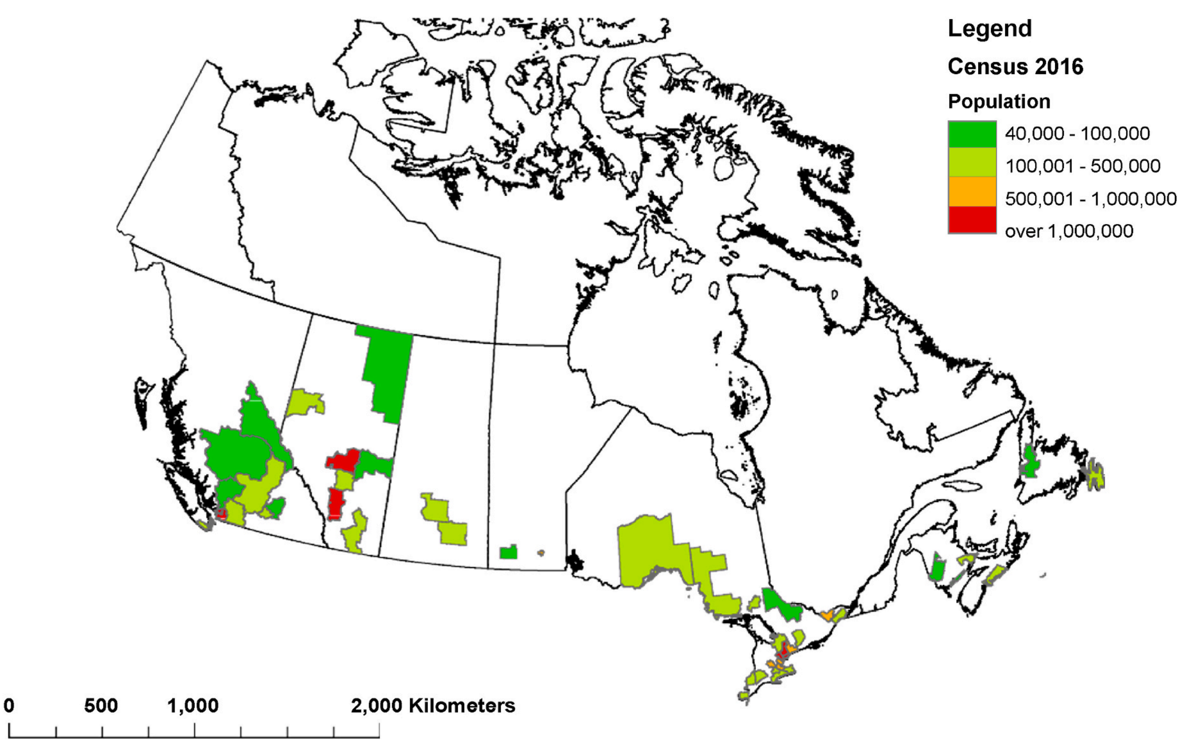

Fig. 1. The 47 study Census Divisions in Canada with population above 40,000.

Surveillance (NAPS) monitors, maintained by Environment and Climate Change Canada (ECCC, https://www.canada.ca/en/environment-climate-change.html). For those CDs with multiple NAPS monitors, average values were calculated. It is worth noting that the 3hMax concentration is a common report concentration for O_3 but not for $PM_{2.5}$ or NO_2 (the WHO guidelines and national standards are given in terms of daily averaged values for $PM_{2.5}$ and NO_2). In this study, we focused on their combined effects, thus the same metric for all of them better represents our study goal (constructing AQHI in Canada) and provides easier interpretation.

In addition to air pollution, daily mean temperature was obtained from the National Climate Data and Information Archive of ECCC for 2001–2018. Like air pollution, for those CDs with multiple weather stations, their averages were used to account for the effect of temperature on health outcomes. Table 1 shows summary statistics over the 47 CDs on average daily counts of mortality and hospitalization, 3hMax concentrations of the three air pollutants, and temperature during study period, while those for each CD are summarized in Supplementary Materials (Tables S1 and S2). Table 1 shows that the seniors account for most of the mortality (all-cause, circulatory, and respiratory), while non-seniors account for all-cause hospitalization. On the other hand, the concentrations of O_3 are much higher than the other two pollutants.

2.2. Statistical analyses

2.2.1. Case-crossover design for city-specific health risks

The case-crossover study design has been widely used, in particular, for short-term exposure to air pollution (see, e.g., Huang et al., 2022; Di et al., 2017; Wing et al., 2017). It is employed in this study to estimate the associations between short-term exposure to air pollutants and health outcomes. We consider a multi-pollutant model to reflect concurrent exposures to $PM_{2.5}$, NO_2, and O_3 in relation to mortality (or hospitalization). For subject i in a given CD, the hazard regression model is:

$$\lambda_i(t) = \Lambda_i(t)\exp\left\{\mathbf{X}(t)^{\mathrm{T}}\boldsymbol{\beta}\right\}, \tag{1}$$

where $\Lambda_i(t)$ is the baseline hazard function for subject i at time t, $\mathbf{X}(t)$ is a vector of covariate of interest, and $\boldsymbol{\beta}$ is a vector of regression coefficients. $\Lambda_i(t)$ is defined as the probability of experiencing the health outcome of interest, when all other covariates equal zero, and the covariates are expected to have effects on the health outcome.

Case-crossover analyses involve choosing control days such that each subject's baseline hazard is similar to that on the case day, $\Lambda_i(t_0) \approx \Lambda_i(t_j)$, where t_0 represents the case day and t_j, $j \in (1, \ldots, J)$ represent the control days (Janes et al., 2005). This assumption requires the control days be close to the case day in many ways. In this study, for the same subject, we use a time-stratified design with control days sharing the same weekdays and calendar month as the case day. Note that the number of control days might be different for different cases. Any other individual-level covariates (e.g., age, sex, and socioeconomic status) are not considered to be confounders in case-crossover models, as they remain constant on both case and control days.

Under this assumption, the hazard ratio takes the form as follows: $\lambda_i(t_0)/\lambda_i(t_j) = \exp[\{\mathbf{X}(t_{i0}) - \mathbf{X}(t_{ij})\}^{\mathrm{T}}\boldsymbol{\beta}]$. For a conditional logistic model for case-crossover data, we denote P_i as the probability that subject i experiences case at time 0 and control times at time $j = 1, \ldots, J$ as follows:

$$P_i = \frac{\exp\left\{\mathbf{X}(t_{i0})^{\mathrm{T}}\boldsymbol{\beta}\right\}}{\sum_{j=0}^{J}\exp\left\{\mathbf{X}(t_{ij})^{\mathrm{T}}\boldsymbol{\beta}\right\}}, \tag{2}$$

where the predictors at time j for subject i, $\mathbf{X}(t_{ij})$, include concentrations of $PM_{2.5}$, NO_2, O_3 and the natural spline functions of temperature. Since people are exposed to multiple pollutants simultaneously and the pollutants are correlated with each other, this 3-pollutant health model is a proximity of the complex multi-pollutant exposures and thus a suitable model used to investigate their health impacts simultaneously. To capture potential non-linear confounding effects of temperature, the spline functions for temperature are generated with 4 degrees of freedom, and the reference comfortable temperature is set to 18 °C.

2.2.2. Hierarchical model for national health risks

A Bayesian hierarchical model is used to combine the health effects across CDs to obtain a national estimate of adverse health effects of air pollution exposures (Huang et al., 2022). For CD k, we denote the pollution and health data as Y_k and H_k, respectively, the coefficients for the three pollutants in model (1) as $\boldsymbol{\beta}_k = \{\beta_{kp}\}$ for $p = 1, 2, 3$ and $k = 1, \ldots, K$ ($K = 47$ in this study), and the national health risk as $\overline{\boldsymbol{\beta}}$. According to (2), we build a hierarchical model:

Table 1
Summary statistics of average daily health counts and exposures across 47 study Census Divisions.

	Cause	Age-group	Mean	Min	1st Qu.	Median	3rd Qu.	Max
Mortality	All-cause	All-age	6.9	0.2	2.1	3.6	8.6	42.8
		Non-senior	1.4	0.1	0.4	0.8	1.7	8.5
		Senior	5.4	0.1	1.7	2.9	6.8	34.3
	Circulatory	All-age	2.2	0.1	0.7	1.2	2.7	13.1
		Non-senior	0.3	0	0.1	0.2	0.4	1.8
		Senior	1.9	0	0.6	1	2.4	11.3
	Respiratory	All-age	0.6	0	0.2	0.3	0.8	3.7
		Non-senior	0.1	0	0	0	0.1	0.3
		Senior	0.6	0	0.2	0.3	0.7	3.4
Hospitalization	All-cause	All-age	72.8	6.2	23.8	39.8	83.5	424
		Non-senior	46.2	4.6	14.3	23.6	50.9	257.5
		Senior	26.6	1	9.9	16.4	31.2	166.5
	Circulatory	All-age	10.8	0.7	4.3	6.8	13.9	67.1
		Non-senior	3.9	0.3	1.5	2.4	5	22.7
		Senior	6.9	0.2	2.7	4.3	8.6	44.4
	Respiratory	All-age	6.9	0.6	2.5	4.2	8.1	38.6
		Non-senior	3.4	0.4	1.3	2.2	4.1	17.2
		Senior	3.5	0.2	1.3	2.1	4	21.5
$PM_{2.5}$ (μgm^{-3})			12.3	9	11.2	12.1	13.5	19.1
NO_2 (parts per billion, ppb)			16.7	7.1	12.9	16.9	19.5	30
O_3 (ppb)			35.8	28.5	33.3	34.8	38.9	43.6
Temperature ($^{\circ}C$)			6.4	0.3	4.4	6.9	8.5	10.5

Note: the quintiles correspond to daily average values of each CD.

$$\begin{aligned}&\left[H_k|Y_k,\ \boldsymbol{\beta}_k\right]\ \sim\ \prod_i P_{ik}\\&\boldsymbol{\beta}_k\ \sim\ N(\overline{\boldsymbol{\beta}},\ \overline{\boldsymbol{\Sigma}}).\end{aligned}\tag{3}$$

For each city we fit the case-crossover design and conditional logistic model to obtain the maximum likelihood estimation of $\boldsymbol{\beta}_k$, denoted by $\widehat{\boldsymbol{\beta}}_k$. By Taylor expansion, we construct a normal approximation to the log likelihood, $\ell(\boldsymbol{\beta}_k) = \log[H_k|Y_k, \boldsymbol{\beta}_k]$, by using a second order expansion as follows (note that $\frac{\partial \ell(\boldsymbol{\beta}_k)}{\partial \boldsymbol{\beta}_k}|_{\boldsymbol{\beta}_k=\widehat{\boldsymbol{\beta}}_k} = 0$):

$$\begin{aligned}\ell(\boldsymbol{\beta}_k)\ &\approx\ \ell(\widehat{\boldsymbol{\beta}}_k) - \frac{1}{2}(\boldsymbol{\beta}_k\text{-}\widehat{\boldsymbol{\beta}}_k)^{\mathrm{T}}\widehat{\boldsymbol{\Sigma}}_k^{-1}(\boldsymbol{\beta}_k\text{-}\widehat{\boldsymbol{\beta}}_k)\\ \widehat{\boldsymbol{\Sigma}}_k^{-1}\ &=\ -\frac{\partial^2\ell(\boldsymbol{\beta}_k)}{\partial\boldsymbol{\beta}_k^2}|_{\boldsymbol{\beta}_k=\widehat{\boldsymbol{\beta}}_k},\end{aligned}\tag{4}$$

where $\widehat{\boldsymbol{\Sigma}}_k^{-1}$ is obtained from the case-crossover design and conditional logistic model as well. This implies that $(\widehat{\boldsymbol{\beta}}_k,\ \widehat{\boldsymbol{\Sigma}}_k)$ is a sufficient statistic for H_k. Therefore, instead of modeling (3), we model (5) which is implemented in Stan (Carpenter et al., 2017) with a prior, $\overline{\boldsymbol{\beta}} \sim \text{Half} - \text{Normal}\left(\mathbf{0}, 0.001^2\mathbf{I}\right)$.

$$\begin{aligned}&\widehat{\boldsymbol{\beta}}_k\ \sim\ N(\boldsymbol{\beta}_k,\ \widehat{\boldsymbol{\Sigma}}_k)\\&\boldsymbol{\beta}_k\ \sim\ N(\overline{\boldsymbol{\beta}},\ \overline{\boldsymbol{\Sigma}}).\end{aligned}\tag{5}$$

This is a weakly informative prior, representing our belief that air pollution can not have protective effects on public health, and the 95th percentile of the magnitude of the effect associated with a 10 unit increase of pollutant is assumed to be about 2% relative increase in daily case rate. We choose this prior based on the estimated air pollution health effects reported in the existing literature. For example, Di et al. (2017) reported that each short-term increase of 10 μgm^{-3} in $PM_{2.5}$ and 10 ppb in warm-season ozone were statistically significantly associated with a relative increase of 1.05% and 0.51% in daily mortality rate, respectively. Liu et al. (2019) reported that an increase of 10 μgm^{-3} in the 2-day moving average of PM_{10} concentration was averagely associated with increases of 0.44% in daily all-cause mortality, 0.36% in daily cardiovascular mortality, and 0.47% in daily respiratory mortality.

For comparison, we also present the results from a non-informative prior $\overline{\boldsymbol{\beta}} \sim \text{Normal}\left(\mathbf{0}, 100^2\mathbf{I}\right)$. This second analysis is equivalent to a conventional random-effects meta-analysis commonly used in air pollution studies (see e.g., Liu et al., 2019), albeit using Bayesian rather than Frequentist inference. A fixed effect meta-analysis is a special case of model (5) with $\overline{\boldsymbol{\Sigma}} = \mathbf{0}$, a model that was ruled out as it produces unreasonably small standard errors. For prior distributions, we decompose $\overline{\boldsymbol{\Sigma}} = \text{diag}(\boldsymbol{\sigma})WW^{\mathrm{T}}\text{diag}(\boldsymbol{\sigma})$, where W is the lower-triangular Cholesky factor of a correlation matrix, and the standard parameter $\sigma \sim \exp(\text{scale} = 0.005)$, following the advice of Simpson et al. (2017) that the standard deviations are encouraged to be small and the data must provide evidence for random effects to stray from zero. We chose the prior $W \sim \text{LKJ}(1)$ which is suggested by the Stan users guide (https://mc-stan.org/). The $\text{LKJ}(\cdot)$ is a method used to generate random correlation matrices, proposed by Lewandowski et al. (2009).

2.2.3. Implementation

For all the analysis in this study, we exclude data on common holidays in Canada to avoid any potential bias, including Good Friday, Easter Sunday, Easter, Easter Monday, Canada Victoria Day, Canada Labour Day, Canada Thanksgiving Day, Canada Civic Provincial Holiday, Canada Day, Christmas and New Year holidays. We have run models with different exposures, including 0-day lag, 1-day lag, and the average of 2-day lag (0–1 day), and found the averaged 2-day lag model generally returned a higher effect than the other two. Therefore, we use a 2-day moving average for pollution concentrations throughout this study, which is a commonly used exposure in the existing literature, such as by Liu et al. (2019); Di et al. (2017); Chen et al. (2017). We present the main results of city-specific in terms of odds ratio, $\text{OR} = \exp(\text{IQR}_p\beta_p)$, while presenting national estimates for each pollutant in terms of odds ratio in percentage (%) calculated as $\text{OR}(\%) = 100\cdot[\exp(\text{IQR}_p\beta_p) - 1]$, where β_p is the log odds ratio of pollutant p, and the IQR (Interquartile range) values were calculated using all the exposure values during 2001–2018 across all 47 CDs (7.6 μgm^{-3}, 12.8 ppb, 14 ppb for $PM_{2.5}$, NO_2, and O_3, respectively). For national estimates, we also show the combined risk from the 3 pollutants, $\text{OR}_{\text{combined}}(\%) = 100\cdot\left[\exp\left(\sum_p 0.33\text{IQR}_p\beta_p\right) - 1\right]$. Since people were exposed to multiple pollutants simultaneously, this combined risk roughly represents the practical risk of exposure to these three pollutants.

Furthermore, apart from the analysis of each city and the national model, we also categorized the 47 CDs into 3 groups based on their population sizes: S-group of 11 small cities (<100,000); M-group of 24 medium cities (100,000 to 500,000) and L-group of 12 large cities (>500,000). This is to see if the most influential pollutant for each outcome is consistent for all sizes or if the national signal is simply dominated by the biggest cities.

The case-crossover model had been implemented using the conditional logistic regression function in *R*, while the hierarchical model for combining city-specific effects had been implemented in Stan (Carpenter et al., 2017) based on *R* interface. The Stan code and *R* script are publicly shared on Github (https://github.com/cghr-toronto/public/tree/master/PollutionHealth47CDs).

3. Results

3.1. City-specific health effects of air pollution

The odds ratios for each city on all-age and -cause mortality and hospitalization associated with an IQR-unit increase in each pollutant are shown in Fig. 2, while those from senior and non-senior all-cause results are presented in Supplementary Materials B to save space. In Fig. 2, Ontario (the most populous province in Canada) CDs are indicated in red to show homogeneity (e.g., $PM_{2.5}$ effects on hospitalization, Fig. 2b) or heterogeneity (e.g., $PM_{2.5}$ effects on mortality, Fig. 2a) among these 21 Ontario CDs. The length of line represents the 95% confidence interval (CI) while the dark vertical line represents the estimated national effect with 95% HDI (highest density interval) as dashed lines. It is shown that the 95% CIs of odds ratios from hospitalization are in general much narrower than those from mortality, which is mainly because of the larger daily counts of hospitalization. Fig. 2 also shows that NO_2 was the main pollutant affecting all-age and -cause mortality and hospitalization, because most of the lines (95% CIs) in Fig. 2c and d have higher middle points (posterior medians) than the neutral effect 1. For both $PM_{2.5}$ and O_3, they presented their important impacts in most of the cities in all-age and -cause mortality, but not for all-age and -cause hospitalization. In addition, we compare the odds ratios (across 47CDs) between mortality and hospitalization on all-age and -cause health outcomes (see Fig. S1 in Supplementary Materials B), and it is found that in cities where the pollution effect on mortality is large (higher OR), the effect on hospitalization is generally also slightly large.

Similarly, the odds ratios for each city on all-age respiratory mortality and hospitalization associated with an IQR-unit increase in each pollutant are shown in Fig. 3, while the circulatory related results are presented in Fig. 4. The former shows that all three pollutants were significantly affecting all-age respiratory hospitalization, since the lower 95% HDI of their national combined effects were obviously higher than neutral effect 1. O_3 was the main pollutant affecting all-age respiratory mortality as well, while the impact of NO_2 was more uncertain because the lower 95% HDI of its national combined effect is very close to the neutral effect.

In Fig. 4, NO_2 was the main pollutant affecting all-age circulatory

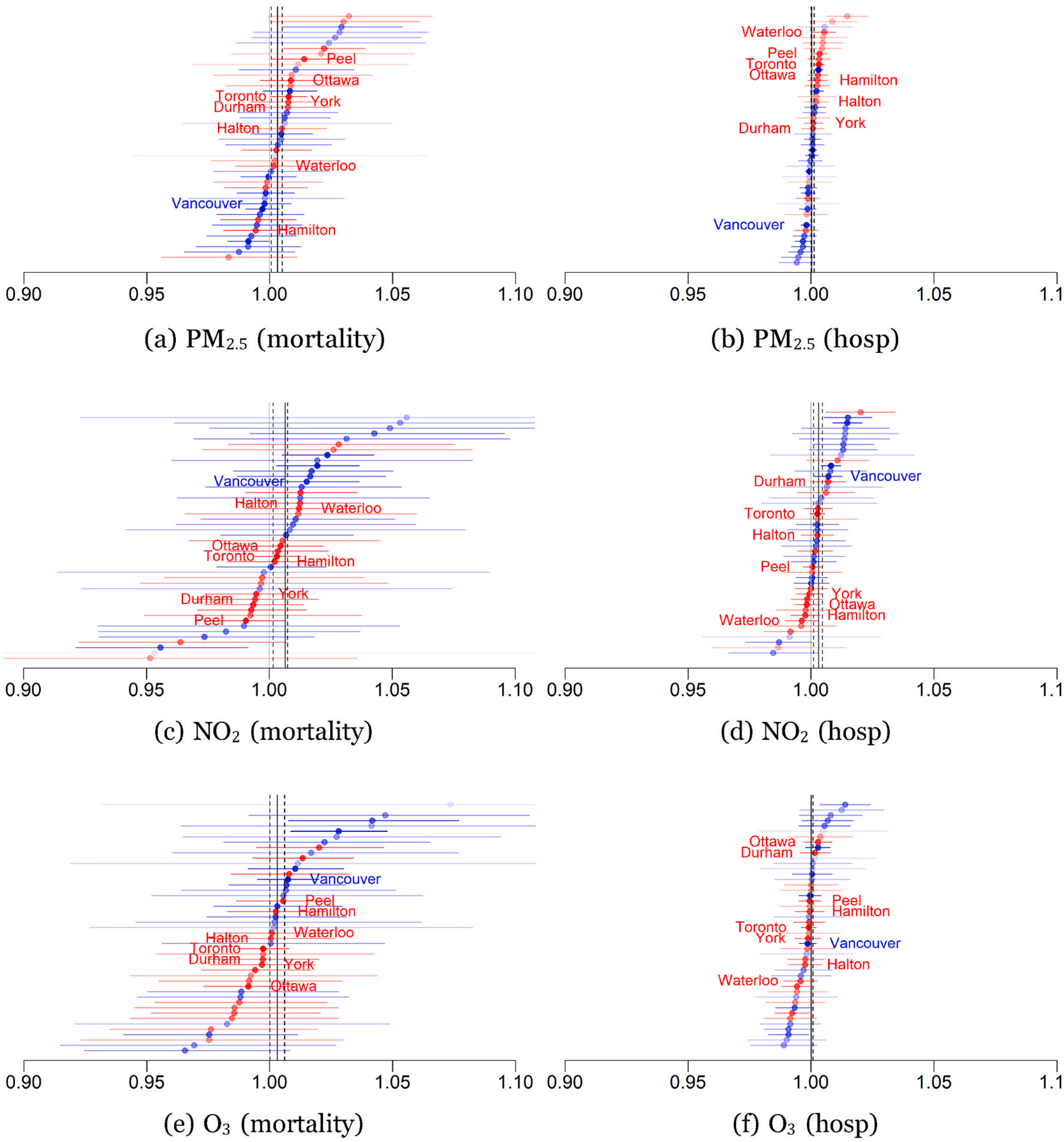

Fig. 2. Odds ratio associated with an IQR-unit increase for each pollutant on all-age and -cause mortality from 2001 to 2015 (left) and hospitalization from 2001 to 2018 (right), with Ontario CDs being in red. The length of line represents the 95% CI of the estimated odds ratio, while the dark vertical line represents estimated national effect with 95% HDI as dashed lines.

mortality. Both $PM_{2.5}$ and O_3 seemed to affect all-age circulatory mortality but with more uncertainty, and this is also the case for the impact of NO_2 on all-age circulatory hospitalization. For the estimated odds ratios on mortality, there is more uncertainty in respiratory compared to circulatory, since those 95% CIs in Fig. 3 (left panel) are generally much wider than those from Fig. 4 (left panel). More health results from different age groups are presented in Supplementary Materials B: Figs. S2–S4 for seniors, and Figs. S5–S7 for non-seniors.

3.2. National health effects of air pollution

For national health effects of air pollution, we consider the excess of odds ratios in percentage (%, calculated as $100 \cdot \left[\exp\left(\mathrm{IQR}_p \beta_p\right) - 1\right]$ for each pollutant, where β_p is the log odds ratio) associated with an IQR-unit increase in pollutant p. Table 2 summarizes the estimated national air pollution effects and their HDI for 6 cause-specific health outcomes attributable to the three air pollutants with each having three age groups, all-age, senior, and non-senior. The last column indicates the combined effects (in %) of the three air pollutants together with per IQR/3 unit increase in each pollutant.

Table 2 clearly shows that O_3 has an effect on respiratory conditions, both mortality and hospitalization. For all-age group, it is likely to have a 0.17%–2.78% increase in the odds of all-age respiratory mortality for every 14 ppb increase of O_3, while it is 0.62%–1.46% for all-age respiratory hospitalization. NO_2 is the main pollutant affecting both all-age and -cause mortality and hospitalization, and it seems that NO_2 affected respiratory hospitalization rather than circulatory hospitalization. For all-age group, it is likely to have 0.57%–1.47% increase in the odds of respiratory hospitalization for every 12.8 ppb increase of NO_2. Table 2 also shows that respiratory hospitalization is notable that all three pollutants have an effect on all-age and non-senior groups, and is the only analysis where $PM_{2.5}$ is significantly important (the lower of

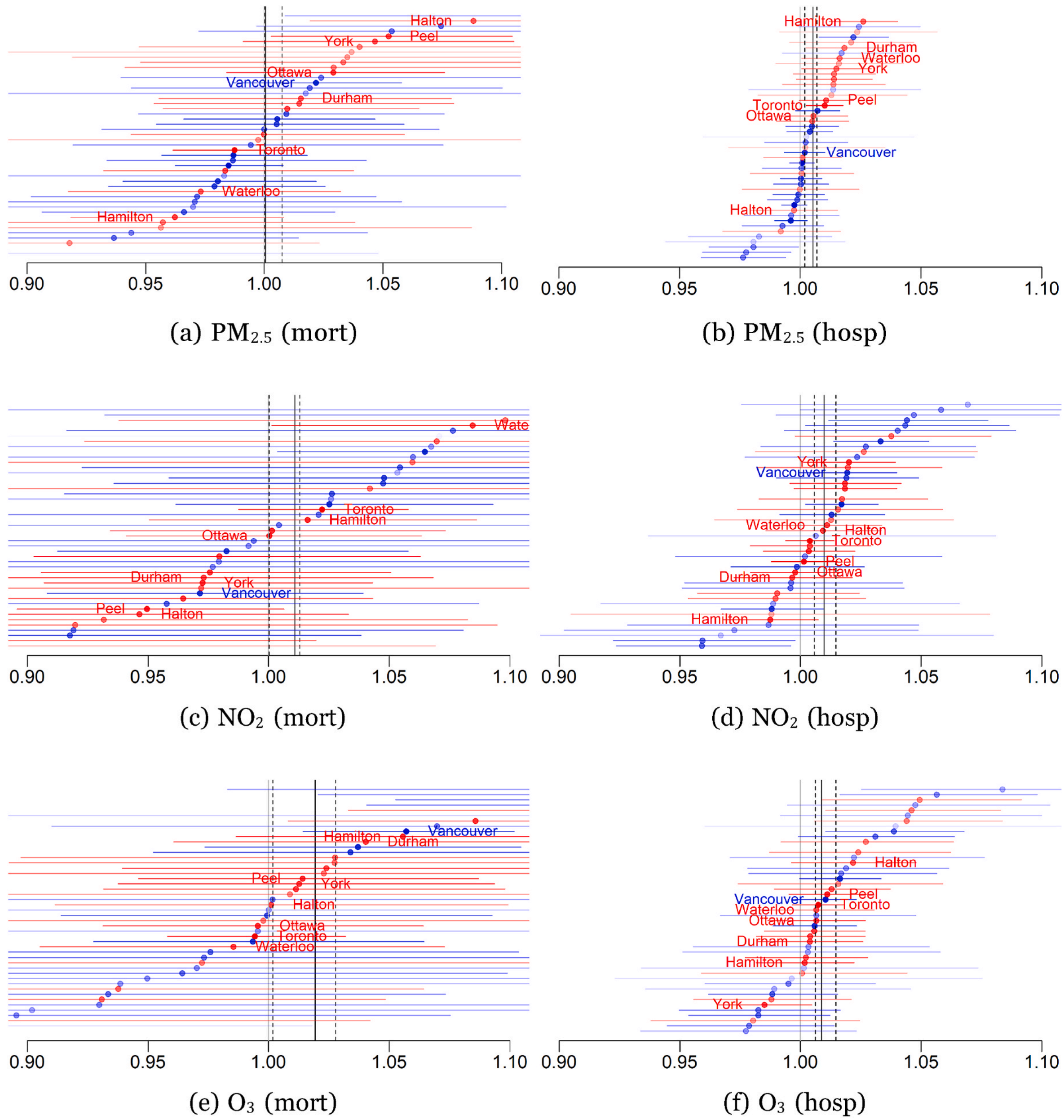

Fig. 3. Odds ratio associated with an IQR-unit increase for each pollutant on all-age respiratory mortality from 2001 to 2015 (left) and hospitalization from 2001 to 2018 (right), with Ontario CDs being in red. The length of line represents the 95% CI of the estimated odds ratio, while the dark vertical line represents estimated national effect with 95% HDI as dashed lines.

95% HDI is much higher than 0%), with 7.6 μgm^{-3} increase of $PM_{2.5}$ causing 0.19%–0.69% (0.33%–1.1%) increase in the odds of all-age (non-senior) respiratory hospitalization. In addition, $PM_{2.5}$ also affected mortality conditions (all-age and -cause, senior all-cause, senior respiratory, all-age circulatory, non-senior circulatory) but with more uncertainty. It is also noted from Table 2 that the seniors account for most of the mortality effects. For example, the estimated OR (%) of $PM_{2.5}$ on all-age and -cause mortality is 0.31%, while it is 0.42% on seniors but 0.01% on non-seniors; the estimated OR (%) of NO_2 on all-age and -cause mortality is 0.64%, while it is 0.91% on seniors but 0.17% on non-seniors. Among the 6 cause-specific health outcomes in Table 2, the pollutant-combined effects of the all-age group are higher for both respiratory mortality and hospitalization where the non-senior groups give higher relative risk than the senior groups.

In addition, we present the national health effects from a non-informative prior $\overline{\boldsymbol{\beta}} \sim \mathrm{Normal}(\mathbf{0}, 100^2\mathbf{I})$ used in (5), shown as Table S3 in Supplementary Materials C. Similar to the results from Table 2, NO_2 is the main pollutant affecting respiratory hospitalization, while O_3 is affecting both respiratory mortality and hospitalization. For $PM_{2.5}$, it has a negligible adverse effect across various health outcomes (except for all-age and senior all-cause mortality, and all-age and non-senior respiratory hospitalization).

National health risks are compared with three population-sized groups: L-group of 12 large cities (>500,000), M-group of 24 medium cities (100,000 to 500,000), and S-group of 11 small cities (<100,000). The comparison results for $PM_{2.5}$, NO_2 and O_3 are presented in Supplementary Materials D (Figs. S9–S11). Overall, the health effect size (point estimate) is comparable among the three groups, whereas the uncertainty in the 95% HDI for the adverse health effects is generally narrower for the L-group but wider for the S-group. The combined effects from the L-group and M-group are generally more consistent with the national combined effects than those from the S-group.

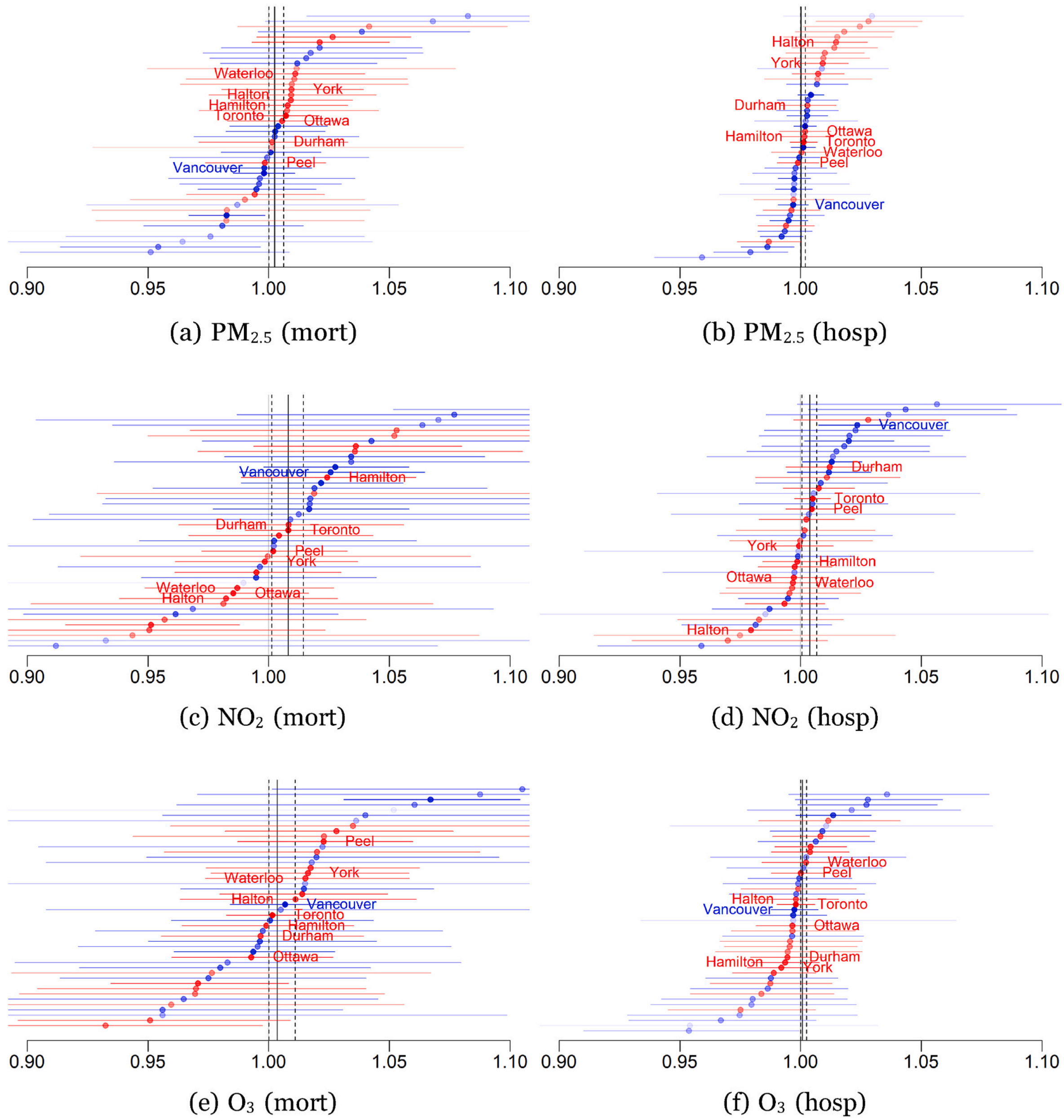

Fig. 4. Odds ratio associated with an IQR-unit increase for each pollutant on all-age circulatory mortality from 2001 to 2015 (left) and hospitalization from 2001 to 2018 (right), with Ontario CDs being in red. The length of line represents the 95% CI of the estimated odds ratio, while the dark vertical line represents estimated national effect with 95% HDI as dashed lines.

4. Discussion

This study investigated the short-term effects of three main air pollutants ($PM_{2.5}$, NO_2, and O_3) on mortality and hospitalization based on 47 Census Divisions in Canada with each having more than 40,000 population, which covers more than 60% of the population in Canada. Due to differing data availability, the study periods were 15 years (2001–2015) for mortality and 18 years (2001–2018) for hospitalization. We employed a case-crossover multi-pollutant model to investigate the short-term effects of the three air pollutants simultaneously. Among the three air pollutants, we found that NO_2 is the main pollutant affecting both all-age and -cause mortality and hospitalization, and it seems that NO_2 affected respiratory hospitalization rather than circulatory hospitalization. O_3 has an effect on respiratory conditions, both mortality and hospitalization. $PM_{2.5}$ has an important effect on all-age and non-senior respiratory hospitalization and a slight effect on mortality conditions.

A number of studies have been conducted worldwide to investigate the impact of air pollution on human health. Our findings are generally consistent with existing research worldwide and in Canada. For example, our finding of significant O_3 effects on mortality is consistent with an earlier study (1996–2012) of four cities in Canada by Huang et al. (2022) and a nationwide study in 272 Chinese cities by Yin et al. (2017). Our findings of $PM_{2.5}$ impact are comparable to those findings from a national Chinese study by R. Chen et al. (2017) where the $PM_{2.5}$ pollution conditions were much more serious. Specifically, we found that 10 μgm^{-3} increase in the 2-day moving average of $PM_{2.5}$ concentrations was associated with a 0.31% (equivalently, 0.24% for a 7.6 μgm^{-3} increase of $PM_{2.5}$) increase of all-age circulatory mortality while it was 0.27% from the study by R. Chen et al. (2017); a 0.41% (equivalently, 0.31% for a 7.6 μgm^{-3} increase of $PM_{2.5}$) increase of all-age and -cause mortality while it was 0.22% from the study by R. Chen et al.

Table 2
National risk (%) with 95% HDI per IQR-unit increase in each pollutant (7.6 μgm^{-3}, 12.8 ppb, 14 ppb for $PM_{2.5}$, NO_2, O_3, respectively).

Outcome	$PM_{2.5}$ [b]	NO_2 [b]	O_3 [b]	Combined [a]
Mortality, all-cause				
All-age	0.31 (0.06, 0.51)	0.64 (0.15, 0.74)	0.31 (0.01, 0.61)	0.42 (0.24, 0.50)
Senior	0.42 (0.03, 0.63)	0.91 (0.16, 1.01)	0.34 (0.00, 0.78)	0.43 (0.27, 0.63)
Non-senior	0.01 (0.00, 0.48)	0.17 (0.00, 0.79)	0.22 (0.01, 1.08)	0.31 (0.07, 0.56)
Hospitalization, all-cause				
All-age	0.03 (0.00, 0.13)	0.30 (0.11, 0.46)	0.00 (0.00, 0.08)	0.13 (0.06, 0.18)
Senior	0.02 (0.00, 0.11)	0.53 (0.22, 0.86)	0.02 (0.00, 0.19)	0.24 (0.11, 0.33)
Non-senior	0.09 (0.01, 0.15)	0.22 (0.01, 0.30)	0.02 (0.00, 0.12)	0.08 (0.04, 0.15)
Mortality, respiratory				
All-age	0.06 (0.00, 0.75)	1.09 (0.02, 1.29)	1.93 (0.17, 2.78)	1.02 (0.39, 1.29)
Senior	0.24 (0.00, 0.91)	0.20 (0.00, 1.37)	1.96 (0.21, 2.71)	0.71 (0.22, 1.28)
Non-senior	0.02 (0.00, 1.08)	0.24 (0.01, 2.39)	1.02 (0.01, 3.10)	0.95 (0.21, 1.68)
Hospitalization, respiratory				
All-age	0.52 (0.19, 0.69)	0.97 (0.57, 1.47)	0.87 (0.62, 1.46)	0.83 (0.65, 1.02)
Senior	0.14 (0.00, 0.39)	0.66 (0.02, 1.13)	1.21 (0.61, 1.74)	0.68 (0.38, 0.88)
Non-senior	0.71 (0.33, 1.10)	1.38 (0.68, 1.86)	0.88 (0.18, 1.59)	0.91 (0.66, 1.26)
Mortality, circulatory				
All-age	0.24 (0.01, 0.62)	0.81 (0.13, 1.44)	0.35 (0.00, 1.10)	0.50 (0.26, 0.81)
Senior	0.09 (0.03, 0.69)	1.21 (0.16, 1.53)	1.06 (0.05, 1.23)	0.79 (0.26, 0.86)
Non-senior	0.17 (0.00, 0.79)	0.50 (0.00, 1.70)	0.46 (0.01, 2.32)	0.61 (0.15, 1.22)
Hospitalization, circulatory				
All-age	0.02 (0.00, 0.19)	0.38 (0.06, 0.66)	0.08 (0.00, 0.24)	0.14 (0.06, 0.29)
Senior	0.02 (0.00, 0.28)	0.55 (0.02, 0.79)	0.06 (0.00, 0.31)	0.23 (0.08, 0.37)
Non-senior	0.01 (0.00, 0.23)	0.10 (0.00, 0.62)	0.06 (0.00, 0.41)	0.12 (0.05, 0.32)

[a] The last column shows $100 \cdot [\exp(\sum p\ 0.33\mathrm{IQR}_p\beta_p) - 1]$ as the combined risk from the 3 pollutants.
[b] Three columns 2–4 show $100 \cdot [\exp(\mathrm{IQR}_p\beta_p) - 1]$ for each pollutant, where β_p is the log odds ratio of pollutant *p*.

(2017). On the other hand, another recent worldwide study by Liu et al. (2019) also reported a 0.41% increase of all-cause mortality associated with a 10 μgm^{-3} increase of 2-day moving $PM_{2.5}$ concentrations for the Chinese cities, however, it was 1.7% (much higher) for Canadian cities.

The case-crossover model we used here is somewhat less commonly used than Poisson time series analysis for studies of short-term effects of air pollution. The relationship between the two methods is well understood, and relates to assumptions made about the baseline hazard functions (Lu and Zeger, 2007). We have chosen the case-crossover model as any individual-level covariates (e.g., age, sex, race) are automatically adjusted for, as are day-of-week effects. Also, any risk factors changing slowly in time are cancelled out (i.e. seasonal effects). The case-crossover analysis is as a consequence simpler than the Poisson time series analysis, as the time trend does not need to be modelled, and is potentially less prone to model misspecification.

When pooling the odds ratios across CD's, we used a half-normal prior distribution for the national-average health effects of the three air pollutants which restricts the effects to be non-negative, which reflects our belief that air pollution should only have the adverse effects (if any) on human health rather than a protective effect. We note that the normal priors, which allow negative values, have been commonly used for pooling odds ratios across cities, however, the use of half-normal priors here is the statistically rigorous way of reflecting our knowledge that air pollution is not protective (although for one or more pollutants in some cities the adverse effect may be negligible). For the purpose of comparison, we have presented the pooling results from using normal priors in Supplementary Materials C which shows similar results to our findings.

In order to further analyze any temporal and spatial variation of the health effects of pollutants, we have run additional analyses, including the results of the three groups of cities (L-group, M-group, and S-group) and the results of all-age and -cause mortality and hospitalization using only the most recent 10 years data (see Fig. S8 in Supplementary Materials B). It is found that the combined effects from the L-group and M-group are generally more consistent with the national combined effects than those from the S-group. By comparing Fig. 2 and S8, we found that the results from the most recent 10 years of data are very similar to those using the whole time period (2001–2015 for mortality and 2001–2018 for hospitalization).

Finally, we acknowledge that controlling for potential confounders is crucial in epidemiological research, and in our study the potential confounding effect of holidays was controlled by restriction (restricted to non-holidays). Although holidays can be included as part of the variables in the adjusted models, we chose to eliminate them to further control for any potential confounding effects. As Canada has relatively few holidays, it is possible that this restriction did not significantly impact the general results.

Funding source

This study was funded under "Addressing Air Pollution Horizontal Initiative" (#810625) of Health Canada, Canada (Principal Investigator, HH Shin).

Author contributions

G. Huang: data curation, methodology, analysis, draft preparation, and revision; P. Brown: methodology, draft preparation, reviewing and editing; and H.H. Shin: conceptualization, proposal, draft preparation, reviewing and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The code and part of data for this study are available on github: https://github.com/cghr-toronto/public/tree/master/PollutionHealth47CDs. Health data cannot be shared due to Privacy Act Canada.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2023.115598.

References

Anderson, H.R., de Leon, A.P., Bland, J.M., Bower, J.S., Strachan, D.P., 1996. Air pollution and daily mortality in London: 1987-92. BMJ 312 (7032), 665–669.

Bell, M., Davis, D., 2001. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. Environ. Health Perspect. 109 (3), 389–394.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., Riddell, A., 2017. Stan: a probabilistic programming language. J. Stat. Software 76 (1), 1–32.

Chen, H., Wang, J., Li, Q., Yagouti, A., Lavigne, E., Foty, R., Burnett, R.T., Villeneuve, P. J., Cakmak, S., Copes, R., 2016. Assessment of the effect of cold and hot temperatures on mortality in Ontario, Canada: a population-based study. CMAJ Open 4 (1), E48–E58.

Chen, R., Yin, P., Meng, X., Liu, C., Wang, L., Xu, X., Ross, J., Tse, L.A., Zhao, Z., Kan, H., Zhou, M., 2017. Fine particulate air pollution and daily mortality: a nationwide analysis in 272 Chinese cities. In: American Journal of Respiratory and Critical Care Medicine, 196.

Di, Q., Dai, L., Wang, Y., Zanobetti, A., Choirat, C., Schwartz, J.D., Dominici, F., 2017. Association of short-term exposure to air pollution with mortality in older adults. JAMA 318 (24), 2446–2456.

Du, X., Chen, R., Meng, X., Liu, C., Niu, Y., Wang, W., Li, S., Kan, H., Zhou, M., 2020. The establishment of national air quality health index in China. Environ. Int. 138, 105594.

Gasparrini, A., et al., 2015. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. Lancet 386 (9991), 369–375.

Haining, R., Li, G., Maheswaran, R., Blangiardo, M., Law, J., Best, N., Richardson, S., 2010. Inference from ecological models: estimating the relative risk of stroke from air pollution exposure using small area data. Spatial Spatio-Temp. Epidemiol. 1 (2–3), 123–131.

Hodgson, T.A., 1970. Short-term effects of air pollution on mortality in New York City. Environ. Sci. Technol. 4 (7), 589–597.

Hoek, G., Fischer, P., van den Brandt, P., Goldbohm, S., Brunekreef, B., 2001. Estimation of long-term average exposure to outdoor air pollution for a cohort study on mortality. J. Expo. Anal. Environ. Epidemiol. 11, 459–469.

Huang, G., Brown, P.E., Fu, S.H., Shin, H.H., 2022. Daily mortality/morbidity and air quality: using multivariate time series with seasonally varying covariances. J. Roy. Stat. Soc.: Series C (Applied Statistics) 71 (1), 148–174.

Janes, H., Sheppard, L., Lumley, T., 2005. Case-crossover analyses of air pollu- tion exposure data: referent selection strategies and their implications for bias. Epidemiology 16 (6), 717–726.

Lawson, A., Choi, J., Cai, B., Hossain, M., Kirby, R., Liu, J., 2012. Bayesian 2-stage space-time mixture modeling with spatial misalignment of the exposure in small area health data. J. Agric. Biol. Environ. Stat. 17 (3), 417–441.

Lewandowski, D., Kurowicka, D., Joe, H., 2009. Generating random correlation matrices based on vines and extended onion method. J. Multivariate Anal. 100 (9), 1989–2001.

Liu, C., et al., 2019. Ambient particulate air pollution and daily mortality in 652 cities. New Engl. J. Med. 381 (8), 705–715.

Lu, Y., Zeger, S.L., 2007. On the equivalence of case-crossover and time series methods in environmental epidemiology. Biostatistics 8 (2), 337–344.

Pope, I.C., Burnett, R., Thun, M., et al., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA, J. Am. Med. Assoc. 287 (9), 1132–1141.

Rushworth, A., Lee, D., Mitchell, R., 2014. A spatio-temporal model for estimating the long-term effects of air pollution on respiratory hospital admissions in Greater London. Spatial Spatio-Temp. Epidemiol. 10, 29–38.

Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., et al., 2017. Penalising model component complexity: a principled, practical approach to constructing priors. Stat. Sci. 32 (1), 1–28.

Stieb, D.M., Burnett, R.T., Smith-Doiron, M., Brion, O., Shin, H.H., Economou, V., 2008. A new multipollutant, no-threshold air quality health index based on short- term associations observed in daily time-series analyses. J. Air Waste Manag. Assoc. 58 (3), 435–450.

WHO, 2021. Ambient (Outdoor) Air Pollution". World Health Organization.

Wing, J.J., Adar, S.D., Sã¡nchez, B.N., Morgenstern, L.B., Smith, M.A., Lisabeth, L.D., 2017. Short-term exposures to ambient air pollution and risk of recurrent ischemic stroke. Environ. Res. 152, 304–307.

Yin, P., Chen, R., Wang, L., Meng, X., Liu, C., Niu, Y., Lin, Z., Liu, Y., Liu, J., Qi, J., You, J., Zhou, M., Kan, H., 2017. Ambient ozone pollution and daily mortality: a nationwide study in 272 Chinese cities. Environ. Health Perspect. 125 (11), 117006.