

Supplemental Digital Content

Long-term exposure to traffic-related air pollution and cardiovascular mortality

Hong Chen,^{1,2} Mark S Goldberg,^{3,4} Richard T Burnett,⁵ Michael Jerrett,⁶ Amanda J Wheeler,⁷ and Paul J Villeneuve^{5,8}

¹ Department of Epidemiology and Biostatistics, McGill University, Montreal, Quebec

² Public Health Ontario, Toronto, Ontario

³ Department of Medicine, McGill University, Montreal, Quebec

⁴ Division of Clinical Epidemiology, McGill University Health Centre, Montreal, Quebec

⁵ Population Studies Division, Health Canada, Ottawa, Ontario

⁶ Division of Environmental Health Sciences, School of Public Health, University of California, Berkeley, California

⁷ Air Health Science Division, Health Canada, Ottawa, Ontario

⁸ Department of Epidemiology and Community Medicine, University of Ottawa, Ottawa, Canada

Corresponding Author:

Hong Chen

Public Health Ontario

480 University Avenue, Suite 300

Toronto, Ontario M5G 1V2

Tel: 647-260-7109

Email: hong.chen@oahpp.ca

Table of Contents

List of eFigures.....	3
List of eTables.....	4
Methods.....	5
Mortality Data.....	5
Distance to Roadways as Exposure Variables.....	5
Temporal Variation of NO ₂ between Different Sampling Periods of Dense Sampling Campaigns.....	6
Validation of Temporal Stability of Land Use Regression Models.....	6
Indirect Adjustment for Unmeasured Smoking.....	7
Results.....	11

List of eFigures

eFigure 1: Trends in observed annual average concentrations of nitrogen dioxides (in ppb) across fixed-site monitors in (A) Hamilton, (B) Toronto, 1982-2004, respectively

eFigure 2: Estimated effects of possible confounding by smoking on the association between exposure to NO₂ and all cardiovascular mortality using (A) approach 1 and (B) approach 2 for Toronto, (C) approach 1 and (D) approach 2 for Hamilton, and (E) approach 1 and (F) approach 2 for Windsor.

eFigure 3: Distributions of estimated concentrations of NO₂ (ppb) at the addresses of subjects' homes at time of entry, derived from land use regression models, by city, The Ontario Tax Cohort Study.

List of eTables

eTable 1: Distributions of estimated annual mean concentrations of NO₂ (ppb) at the addresses of subjects' homes at time of entry in Toronto across three time periods, according to the two different back-extrapolation methods, The Ontario Tax File Cohort Study, 1982-2004

eTable 2: Prevalence of smoking status, according to five different levels of concentrations of NO₂ among the participants, 35-85 years of age, from the Canadian Community Health Survey in 2001

eTable 3: Distribution of measured body mass index (in kg/m²) ^a according to the quintiles of concentrations of NO₂ among the participants aged 35 years and above, from the Canadian Community Health Survey in 2001

eTable 4: Estimated average concentration of NO₂ using data from fixed-site monitors during the time period of dense sampling campaigns and across the whole years, by city and year

eTable 5: Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and the estimated concentrations of NO₂ from land use regression models (LUR) as well as proximity to major roads and highways, The Ontario Tax Cohort Study, 1982-2004

eTable 6: Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, among study participants in Toronto, The Ontario Tax File Cohort Study, 1982-2004

eTable 7: Rate ratios (RR_{5ppb}) for an increase of 5 ppb in NO₂ and associated 95% confidence intervals (CI) in Toronto for the association between mortality from cardiovascular disease and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, The Ontario Tax Cohort Study, 1982-2004.

eTable 8: Correlations between estimated annual mean concentrations of NO₂ (ppb) across three time periods in Toronto

METHODS

Mortality Data

Vital status was ascertained for all subjects over the follow-up period of 1982-2004 through a probabilistic record linkage to the Canadian Mortality Database. This database provides data on all deaths of Canadians that occurred in Canada as well as most of those that occurred in approximately 20 U.S. states.¹ The cohort was linked according to first, middle, and family names, sex, date of birth, place of residence, and in some cases social insurance number.² Previous work suggests that under-coverage of the deaths is minimal and the accuracy of identifying deaths is around 98%.² Underlying cause of death until 2000 were coded to the ninth revision of the International Classification of Diseases (ICD-9) and afterwards ICD-10 was used. Coronary heart disease was coded as ICD-9: 410-414 or ICD-10: I20-I25, cerebrovascular disease as ICD-9: 430-438 or ICD-10: I60-I69, and all cardiovascular diseases combined as ICD-9: 400-440 or ICD-10: I00-I99. We assumed that those for whom a link to the mortality database was not found were alive at the end of follow-up.

The accuracy of coding cardiovascular diseases as an underlying cause of death on the Canadian death certificates is, however, not known. Previous studies showed that average false positive rates and false negative rates in coding acute myocardial infarction on death certificates in three Canadian provinces using 1984 data were about 5% and 1%, respectively^{3,4}. In studies conducted in the U.S., death certificates showed overestimates of 7%-10% for cardiovascular diseases and 7%-20% for coronary heart disease.⁵⁻⁸ As a result, some misclassification in outcomes is likely but we expect that it should be independent of exposure to air pollution and therefore no differential bias should be introduced into our risk estimates.

Distance to Roadways as Exposure Variable

First, we calculated distances between subjects' postal-code addresses at the time of entry and major traffic roads (primary urban roads, arterial roads). Distances were categorized as

0-50 m, 51-100 m, 101-200 m, 201-300 m, and greater than 300 m.^{9,10} Second, in several previous studies,¹¹⁻¹⁶ the positive associations between traffic proximity and all-cause and circulatory mortality were largely confined to the distance category of living within 50 meters of a major road or within 100 meters from a highway. As a result, we applied the same definition to create a dichotomous exposure variable coded as 1 if living within 50 m of a major road or 100 m of a highway and 0 otherwise.

Temporal Variation of NO₂ between Different Sampling Periods of Dense Sampling Campaigns

To evaluate the temporal variation, we estimated the average concentrations of NO₂ using measurements from fixed-site monitoring stations in the three cities corresponding to the two-week periods during which the dense sampling campaigns were conducted. In addition, we computed the annual mean concentrations of NO₂ in 2002 and 2004, respectively, and an average between the two years.

Validation of Temporal Stability of Land Use Regression Models

Thorough investigation of the temporal stability of land use regression was conducted for Toronto. In doing this, we back-extrapolated the land use regression models to each year between 1982 and 2002 using the methods described previously.¹⁷ We estimated the pairwise correlation between the historically extrapolated estimates of NO₂ and the estimates of NO₂ from the original land use regression models that were developed for 2002 and 2004 at 5000 random sites in Toronto. Pearson's correlation coefficients were fairly stable (varied from 0.85 to 0.95, depending on the year), suggesting that variability in the concentrations of NO₂ in Toronto is primarily spatial in nature and not temporal.

Using the land use regression models of NO₂ between 1982 and 2004, we further estimated the total variance of NO₂ across all postal-code addresses in Toronto and throughout the follow-up period of 23 years. In addition, we estimated in Toronto the variance of NO₂ that was due to temporal variability from 1982 to 2004. This was done by calculating mean exposure for all postal-code addresses each year in Toronto and then variance of the annual

averages over time. The total variance was 23.5 ppb² while the temporal variance was 7.0 ppb². Thus, 70% of the total variation in the concentrations of NO₂ in Toronto between 1982 and 2004 is associated with spatial variability and only 30% with variation over time. This finding is reinforced by the fact that the annual mean concentrations of NO₂ and their rank ordering at fixed-site monitors in Toronto were relatively constant during the follow-up period of 23 years.

We therefore expected that for Toronto the spatial contrast in NO₂ estimated using the land-use regression models provided reasonable estimates of longer-term spatial exposures to traffic-related air pollutants. The representativeness of land use regression models for long-term exposure has also been reported by Su et al (2009), where the authors showed based on 42 fixed-site monitoring stations in Los Angeles, California and breaking the year into a series of two-week averages, that intraclass correlation coefficient for 2 or 3 rounds of measurements of NO₂ was more than 0.9, indicating that the vast majority of the variation in the concentrations of NO₂ was between sites, not within site.¹⁸ Similarly, other epidemiological studies of long-term health effects of air pollution have reported long-term stability in the spatial patterns of ambient concentrations of NO₂; for example in Montreal, Canada,¹⁹ in California, U.S.A.,²⁰ in North Rhine Westphalia, Germany²¹ and in the Netherlands.²²

For Hamilton and Windsor, we would also expect that similar to other Canadian cities, the land use regression models developed in 2002 and 2004 still reflected the spatial pattern of NO₂ in the 1980s and are valid to determine the association between pollution pattern and cardiovascular mortality. However, because few fixed-site monitors existed in Hamilton and Windsor, we were unable to extrapolate the land use regression models back in time, and thus to examine adequately the temporal stability of the land use regressions for the two cities. This is a limitation of this study, and following the reviewer's comment, we have acknowledged this in our revision.

Indirect Adjustment for Unmeasured Smoking

We did not have access to information on some important individual risk factors of cardiovascular diseases such as smoking, obesity, high blood pressure, and diabetes. Smoking is one of the strongest risk factors for cardiovascular disease, because of the strength of its effect and its high prevalence, and thus has the potential to confound. We thus made use of indirect methods to estimate the bias that may have occurred because smoking was not included directly in the models. These "indirect" methods were developed by Axelson²³ and others²⁴. Specifically, we made use of a Monte Carlo sensitivity analysis proposed by Steenland and Greenland (2004),²⁵ and we extended the method to handle continuously measured exposure variables.²⁶

The method allows for the estimation of a "bias factor" which under the null hypothesis of no association between air pollution and cardiovascular mortality implies that the rate ratio would be estimated solely due to the confounding effects of smoking (RR_{bias}).²³ To compute this, we classified the distribution of exposure in a city into quintiles and then we estimated RR_{bias} for subjects in each level of exposure using the following equation (for the lowest quintile RR_{bias} was set to unity):

$$RR_{\text{bias}} = I_{E+}/I_{E-} = \frac{I_o \times \left(1 + \sum_{i=1}^k P_{c,i} \times (RR_{c,i} - 1) \right)}{I_o \times \left(1 + \sum_{i=1}^k P_{g,i} \times (RR_{c,i} - 1) \right)} \quad (1)$$

where I_{E-} refers to the incidence rate of cardiovascular mortality among unexposed subjects (i.e., the lowest quintile of NO_2), I_{E+} denotes the incidence rate of cardiovascular mortality among the subjects in the cohort who are classified as being exposed (i.e., a higher quintile of NO_2), I_o is the incidence rate among those who do not smoke, $P_{c,i}$ and $P_{g,i}$ represent the prevalence of smokers, at level i (e.g., current smoker, $i=1$; ex-smoker, $i=2$), in the exposed cohorts and the unexposed cohort, respectively. $RR_{c,i}$ is the relative risk for cardiovascular-related mortality for smoking at level i .

To estimate area-specific prevalence of smoking, we made use of the Canadian Community Health Survey conducted in 2001 which is a national probability sample of all households in Canada.²⁷ For this analysis, we included the participants from the Canadian Community Health Survey who lived in Toronto, Hamilton, and Windsor at the time of survey. Because

the relationship between the prevalence of smoking and exposure to NO₂ may be influenced by age, sex, income, and other variables, failure to account for these covariables may result in over-estimating the “bias factor”, because many smoking predictors (age, sex, income, etc) are already in the survival model. As a result, we used two approaches based on different scenarios to calculate $P_{g,i}$ and $P_{g,i}$: (1) The proportions of never, current, and former smokers were calculated for each quintile of NO₂ among the participants in a city. The exposure to NO₂ was derived from the land use regression models (see main text for details) and assigned to the participants at their six-character postal code addresses at time of survey. (2) The proportions of never, current, and former smokers were calculated for each quintile of residuals from a model in which the concentrations of NO₂ at the postal-code addresses of the participants were regressed on age, sex, marital status, income, and ecological variables (same predictors used in the fully-adjusted survival models). Using these two approaches, we bounded the “bias factor” using the approach not adjusting the relationship between smoking and NO₂ for any other covariates as one bound, and adjusting for all other variables in the survival model as the other bound. The true bias factor should be somewhere between the bounds. We repeated the analysis for each of the three cities.

Also for Equation 1, we made use of rate ratios for the cardiovascular diseases of interest for current and former smokers using estimates from the American Cancer Society Cancer Prevention Study II, for the follow-up period 1982 until 1988.²⁸ These estimates were rate ratios that were adjusted for age, sex, marital status, and other risk factors in the original study. Because these rate ratios for smoking and cardiovascular mortality varied by age and sex, we re-weighted the age- and sex-specific rate ratios by the underlying age and sex structure of participants of the Canadian Community Health Survey.

The classic formula of indirect adjustment for unmeasured smoking (Equation 1) allowed us to estimate a RR_{bias} for each quintile of NO₂. Because concentrations of NO₂ are on a continuous scale, we required a bias factor per each increase of 5ppb of NO₂ ($RR_{bias-5ppb}$). To do this, we made use of a method developed by Villeneuve et al (2011).²⁶ Specifically, we derived a simple linear regression model with the dependent variable equal to the estimated RR_{bias} for each quintile of concentrations of NO₂ through using Equation 1. The independent variable was the concentration of NO₂ that was sampled randomly from a

uniform distribution of NO₂ for each of the quintile groups. The slope (RR_{bias-5ppb}) obtained from fitting the linear regression model represented the estimate of the amount of confounding by smoking for each 5 ppb increase of NO₂.

Rate ratios for an increase of 5 ppb of exposure to NO₂ that were indirectly adjusted for smoking (RR_{Indirect adj-5ppb}) are computed as:

$$RR_{\text{Indirect-adj-5ppb}} = RR_{\text{Cox model-5ppb}} / RR_{\text{bias-5ppb}} \quad (2)$$

where RR_{Cox-model-5ppb} is the rate ratio for an increase of NO₂ of 5 ppb adjusted for all variables included in the main analysis.

To compute the statistical uncertainty of RR_{Indirect-adj-5ppb}, we used Monte Carlo sampling (100,000 replications) to repeatedly sample from the priors of the prevalence of current and former smokers in each exposure group as well as the rate ratio for the effect of smoking on cardiovascular mortality. Our prior distribution for the prevalence of smoking was computed from a bivariate normal distribution with means equal to the logit of the proportions of current and former smokers in each exposure group. For the rate ratios relating current and former smoking to cardiovascular mortality, we specified a normal distribution with a mean value equal to the natural logarithm of the rate ratio and standard deviations equal to the standard errors from the American Cancer Society Cancer Prevention Study II²⁸. For each replicate, we also re-sampled the observed rate ratio for the association with concentrations of NO₂ (RR_{Cox model-5ppb}) from its estimated normal distribution with the mean and variance estimated using data from the tax cohort. We repeated the Monte Carlo sensitivity analysis for each of the three causes of death and for the study population in each of the three cities.

Similarly, obesity may also be a possible confounding variable. We obtained the distribution of body mass index (weight(kg)/height(m)²) from the 2001 Canadian Community Health Survey. We assessed whether the distributions of body mass index differed between the exposure groups. The analyses were repeated for each of the three cities.

RESULTS

eFigure 1 shows the secular trends of annual mean concentrations of NO₂ at fixed-site monitors in the three cities between 1982 and 2004. The annual mean concentrations of NO₂ and the rank ordering of fixed-site monitors suggested that the spatial distributions of NO₂ did not change appreciably over the follow-up period.

eTable 1 shows the various estimates of concentrations of NO₂ in Toronto, including the original land-use regression models from 2002 and 2004 as well as the back-extrapolated ones. The estimates of concentrations of NO₂ and their variability across the follow-up period were similar: the mean concentration of NO₂ in the first 10 years follow-up (1982-1992) was approximately 24 ppb (inter-quartile range (IQR): 4.7 ppb) and the mean concentration of NO₂ in the entire study period was about 23 ppb (IQR: 4.4 ppb).

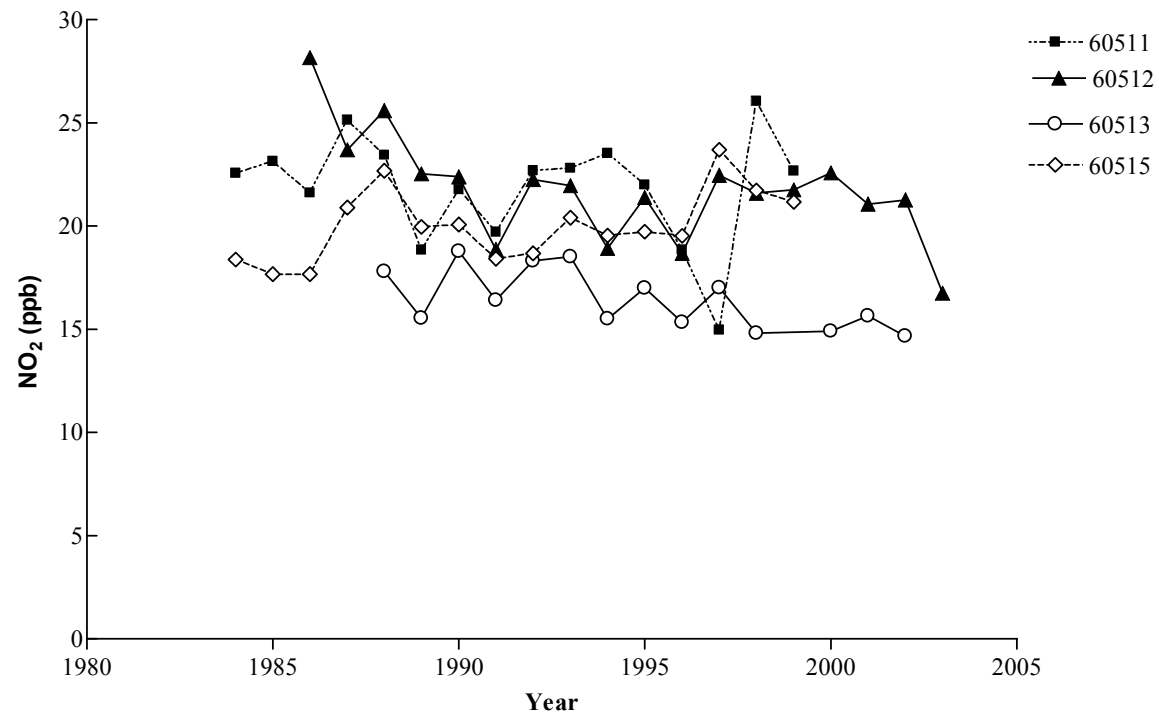
eFigure 2 shows the estimated confounding effects of smoking on the association between exposure to NO₂ and all cardiovascular mortality using two approaches for each of the three cities. The possible confounding effects of smoking appeared to be marginal. This is not surprising, given that the prevalence of current and former smokers were weakly correlated with ambient concentrations of NO₂ in the three cities (**eTable 2**). BMI was found similar across different levels of NO₂ (**eTables 3**), thus BMI was not likely a confounder in this study.

As shown in **eTable 4**, the two-week average concentration of NO₂ during the dense sampling campaign in Hamilton is virtually identical to the annual average concentrations in 2002. There was a marginal difference between the annual mean concentration of NO₂ in 2002 and the two year average of 2002 and 2004. For Toronto, the average concentration of NO₂ during the two dense sampling campaigns was similar to the average concentration of the two years (20.7 ppb versus 20.2 ppb). For Windsor, there was little variation of annual concentration of NO₂ from 2002 to 2004. As a result, we would expect that the estimates from land use regression models developed for the three cities in 2002-2004 capture their long-term average concentrations of NO₂.

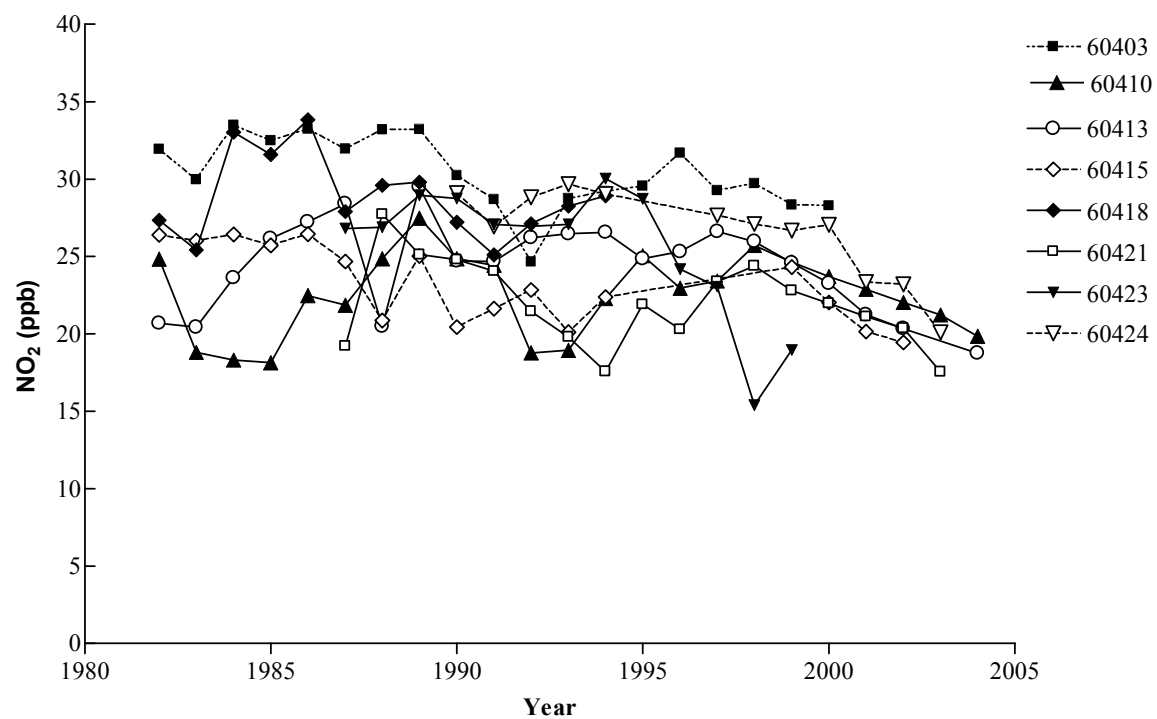
Using the second definition, we estimated rate ratios for the association between cause-specific cardiovascular mortality and living within 50 meters from a major road or 100 meters from a highway (**eTable 5**). We found elevated mortality rates from all cardiovascular diseases and from ischemic heart disease in relation to living in close proximity to a major road or highway.

As shown in **eTable 6**, when the rate ratios were computed for an increase of IQR of NO₂ to account for differences in the absolute value of the distribution, the estimates were similar. **eTable 7** shows that the analyses using historically extrapolated concentrations of NO₂ yielded similar associations to those using the land-use regression data for 2002-2004 in Toronto. For example, the association between ischemic heart disease and the estimates of NO₂ in 1982 is $RR_{5ppb}=1.05$ (95% CI: 1.00-1.10) compared to $RR_{5ppb}=1.06$ (95% CI: 1.00-1.14) for NO₂ in 1992 and $RR_{5ppb}=1.06$ (95% CI: 1.00-1.13) for NO₂ in 2002-2004, suggesting temporal stability of spatial patterns in NO₂ levels in Toronto.

eTable 8 shows correlations between estimated annual mean concentrations of NO₂ (ppb) across three time periods in Toronto, according to the surface maps of NO₂ produced using the land use regression model for 2002-2004 and the two extrapolation methods (based on 5,000 random locations in Toronto).

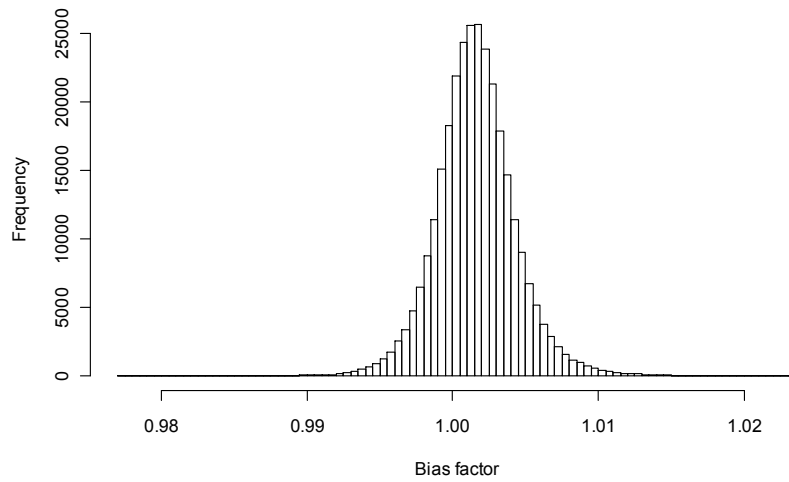


(A) Hamilton

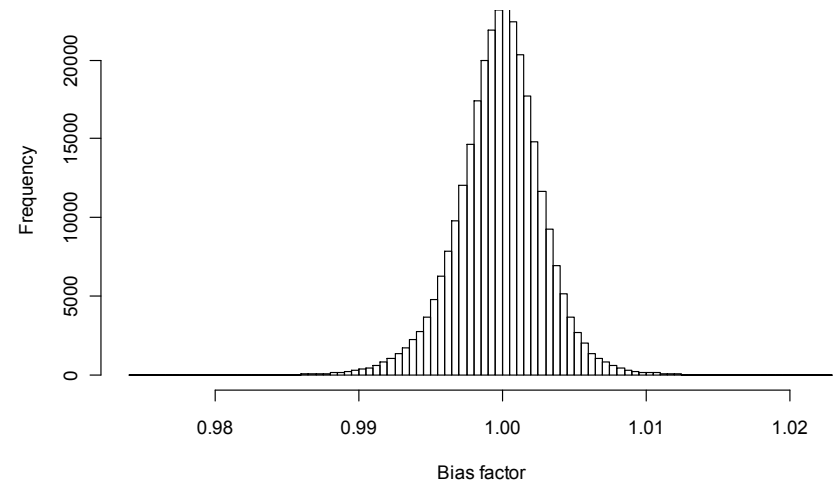


(B) Toronto

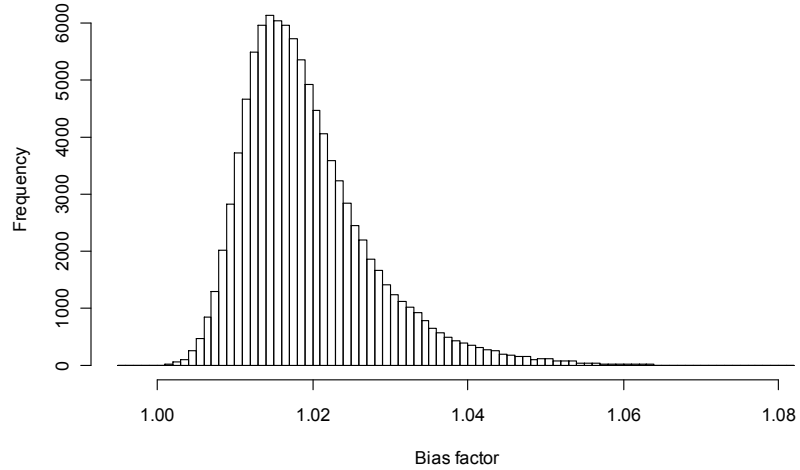
eFigure 1. Trends in observed annual average concentrations of nitrogen dioxides (in ppb) across fixed-site monitors in (A) Hamilton, (B) Toronto, 1982-2004, respectively. The fixed-site monitors are administered by the National Air Pollution Surveillance (NAPS) network in the Ontario region (site number is provided in the legend). For each city, the fixed-site monitors that operated for less than half of the period (≤ 12 years) are not shown. Windsor is not included because only one fixed-site monitoring station was available in the city.



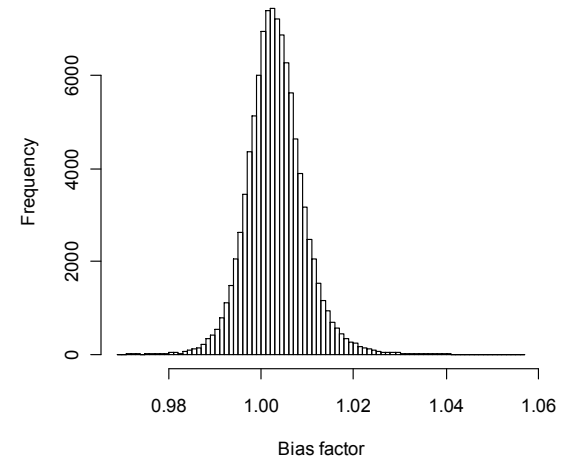
(A) Bias factor by approach 1 (Toronto)



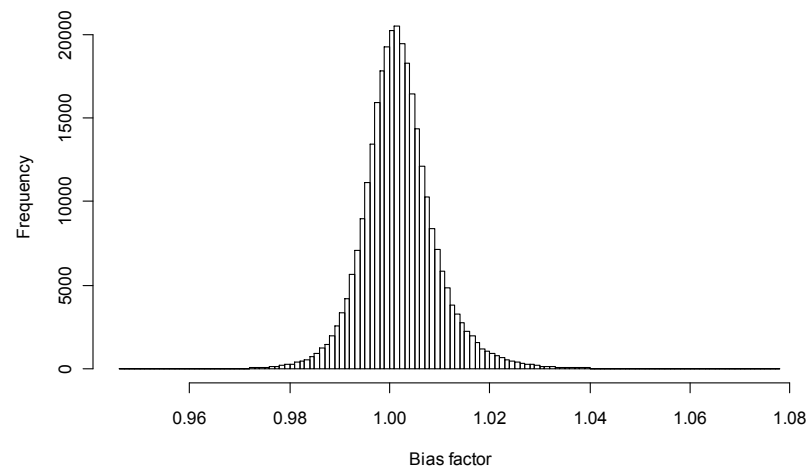
(B) Bias factor by approach 2 (Toronto)



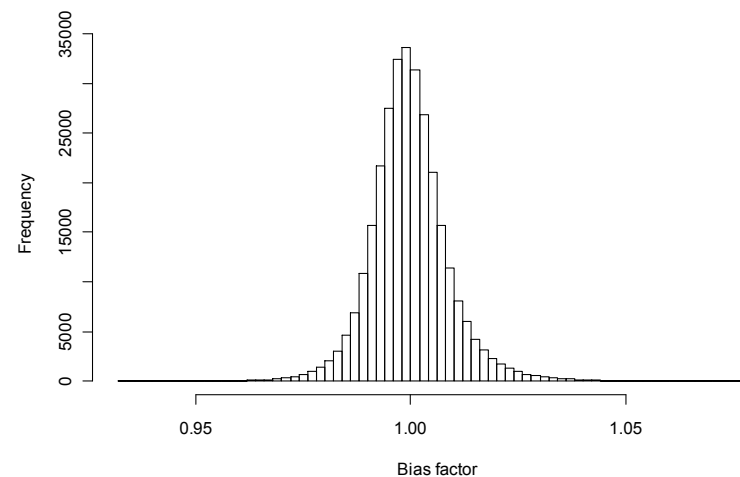
(C) Bias factor by approach 1 (Hamilton)



(D) Bias factor by approach 2 (Hamilton)

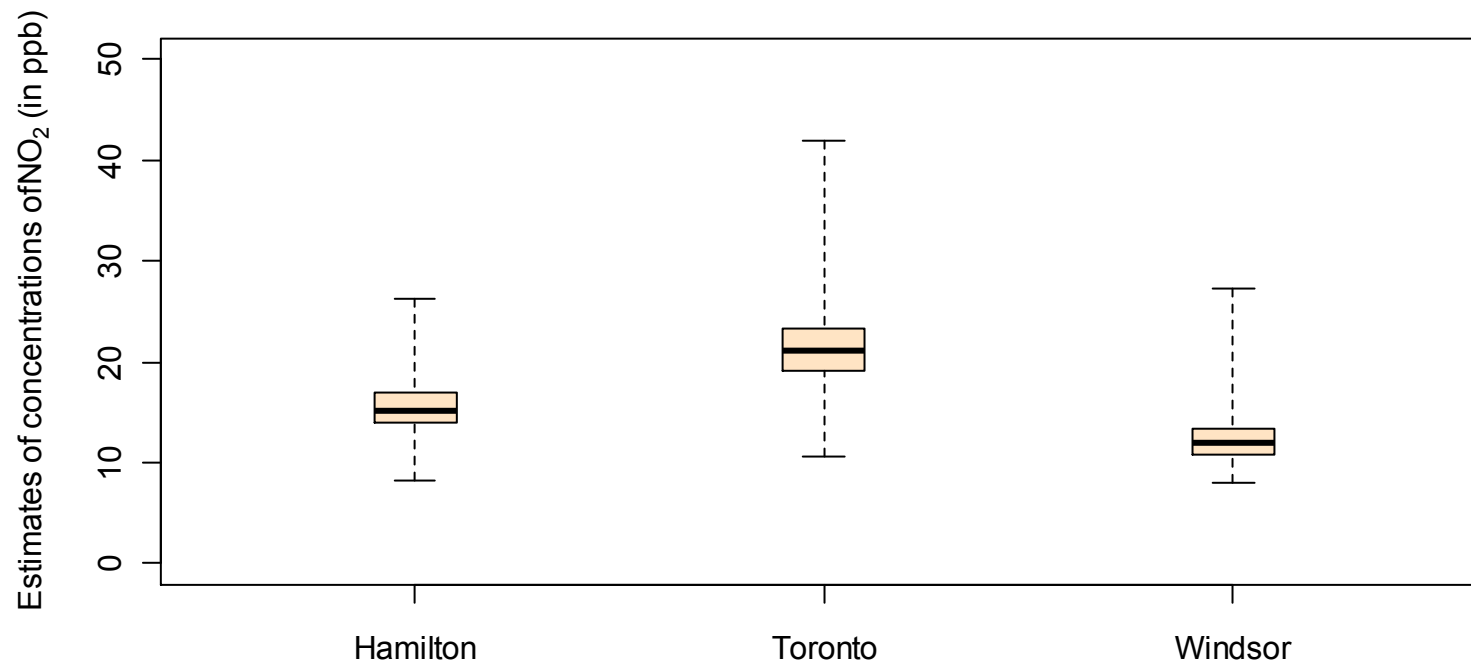


(E) Bias factor by approach 1 (Windsor)



(F) Bias factor by approach 2 (Windsor)

eFigure 2. Estimated effects of possible confounding by smoking on the association between exposure to NO_2 and all cardiovascular mortality using (A) approach 1 and (B) approach 2 for Toronto, (C) approach 1 and (D) approach 2 for Hamilton, and (E) approach 1 and (F) approach 2 for Windsor. The bias factors were estimated using data from 2001 Canadian Community Health Survey.



eFigure 3. Distributions of estimated concentrations of NO₂ (ppb) at the addresses of subjects' homes at time of entry, derived from land use regression models, by city, The Ontario Tax Cohort Study, 1982-2004. For Hamilton, the concentrations of NO₂ were derived from a land use regression model using measurement from a monitoring campaign with 107 monitors in fall 2002.²⁹ For Toronto, the concentrations of NO₂ derived from averaged estimates from two land use regression models using measurements from two monitoring campaigns with 100 monitors in fall 2002 and spring 2004.¹⁵ For Windsor, the concentrations of NO₂ derived from a land use regression model using measurements from four monitoring campaigns with 54 monitors in spring, summer, fall, and winter 2004.³⁰

eTable 1. Distributions of estimated annual mean concentrations of NO₂ (ppb) at the addresses of subjects' homes at time of entry in Toronto across three time periods, according to the two different back-extrapolation methods, The Ontario Tax File Cohort Study, 1982-2004

Exposure metrics	Year	Mean	Minimum	25 th percentile	Median	75 th percentile	Maximum	Interquartile Range
1. Original LUR model ^a	Mean of 2002-2004	21.68	10.61	19.15	21.05	23.24	42.03	4.09
2. IDW-based extrapolation ^b	1982	25.25	12.37	22.65	25.16	27.47	49.78	4.82
	1992	25.24	12.07	22.53	25.17	27.53	47.54	5.00
	Mean of 1982-1992	25.25	12.22	22.67	25.13	27.44	48.12	4.77
3. LUR-based extrapolation ^b	1982	24.44	11.96	21.13	23.70	26.64	55.37	5.51
	1992	23.67	12.52	21.01	23.19	25.69	41.11	4.68
	Mean of 1982-1992	24.06	12.92	21.33	23.55	26.00	47.24	4.67
	Mean of 1982-2004	23.26	12.49	20.62	22.71	25.02	45.50	4.40

^a. Concentrations of NO₂ derived from the average of two land use regression models using measurements from 2002 fall and 2004 spring monitoring campaigns with 100 monitors.^{16,32}

^b. IDW, inverse distance weighted interpolation; LUR, land use regression model

eTable 2. Prevalence of smoking status, according to five different levels of concentrations of NO₂ among the participants, 35-85 years of age, from the Canadian Community Health Survey in 2001

Exposure metrics	Hamilton (n=800)			Toronto (n=1430)			Windsor (n=469)		
	% Never smoker	% Current smoker	% Former smoker	% Never smoker	% Current smoker	% Former smoker	% Never smoker	% Current smoker	% Former smoker
<i>NO₂ (ppb) ^a</i>									
1st quintile ^b	31.2	27.2	41.6	39.2	24.7	36.2	33.3	22.7	44.0
2nd quintile	34.1	29.4	36.5	46.8	20.4	32.8	27.6	33.3	39.1
3rd quintile	29.4	28.6	42.1	45.8	20.3	33.9	40.0	22.9	37.1
4th quintile	30.2	29.4	40.5	44.3	19.6	36.2	27.5	31.8	40.7
5th quintile	28.8	30.4	40.8	38.7	23.0	38.3	40.1	26.1	33.8

^a. For Hamilton, the levels of concentrations of NO₂ were derived using residuals from a model in which concentrations of NO₂ at the six-character postal code addresses of participants of the 2001 Canadian Community Health Survey were regressed against age, sex, marital status, income, and ecological variables (same predictors used in the fully-adjusted survival models). The reason for using “adjusted” NO₂ in replace of raw NO₂ here is to prevent over-adjusting for smoking effects on the association between NO₂ and cardiovascular outcomes, because many smoking predictors (age, sex, income, etc) are already in the survival model. Similarly, the “adjusted” NO₂ were used to assess correlations of NO₂ and smoking in Toronto and in Windsor.

^b. For Hamilton, the quintiles of NO₂ (ppb) are: ≤ -1.6 ; $-1.6-(-0.4)$; $-0.4-0.4$; $0.4-1.6$; and ≥ 1.6 . For Toronto, the quintiles of NO₂ (ppb) are: ≤ -2.9 ; $-2.9-(-1.1)$; $-1.1-0.3$; $0.3-2.1$; and ≥ 2.1 . For Windsor, the quintiles of NO₂ (ppb) are: ≤ 1.3 ; $-1.3-(-0.4)$; $-0.4-0.5$; $0.5-1.3$; and ≥ 1.3 .

eTable 3. Distribution of measured body mass index (in kg/m²) ^a according to the quintiles of concentrations of NO₂ among the participants aged 35 years and above, from the Canadian Community Health Survey in 2001

Exposure metrics	Hamilton (n=800)				Toronto (n=1430)				Windsor (n=469)			
	Under weight (%) ^b	Normal (%) ^b	Over weight (%) ^b	Obesity (%) ^b	Under weight (%)	Normal (%)	Over weight (%)	Obesity (%)	Under weight (%)	Normal (%)	Over weight (%)	Obesity (%)
<i>NO₂ (ppb) ^c</i>												
1st quintile ^d	2.1	44.2	36.3	17.4	2.4	53.1	32.3	12.2	1.9	36.5	40.4	21.2
2nd quintile	2.3	39.1	39.1	19.5	2.4	48.2	35.4	14.0	0	38.5	36.5	25.0
3rd quintile	1.2	35.3	43.5	20.0	2.3	49.1	34.3	14.3	0	46.8	29.8	23.4
4th quintile	2.0	36.3	34.3	27.4	5.6	47.2	34.6	12.6	2.6	43.6	41.0	12.8
5th quintile	2.1	47.4	34.7	15.8	2.3	45.7	37.7	14.3	0	51.9	30.8	17.3

^a. The body-mass index is the weight in kilograms divided by the square of the height in meters.

^b. BMI is classified as follows: underweight: < 18.5; normal weight: 18.5-24.9; overweight: 25.0-29.9; obesity: > 30.³¹

^c. For Hamilton, the levels of concentrations of NO₂ were derived using residuals from a model in which concentrations of NO₂ at the six-character postal code addresses of the participants of the 2001 Canadian Community Health Survey were regressed against age, sex, marital status, income, and ecological variables (same predictors used in the fully-adjusted survival models). The reason for using “adjusted” NO₂ in replace of raw NO₂ here is to prevent over-adjusting for smoking effects on the association between NO₂ and cardiovascular outcomes, because many smoking predictors (age, sex, income, etc) are already in the survival model. Similarly, the residuals of NO₂ were used to assess correlations of NO₂ and smoking in Toronto and Windsor.

^d. For Hamilton, the quintiles of NO₂ (ppb) are: ≤-1.6; -1.6-(-0.4); -0.4-0.4; 0.4-1.6; and ≥1.6. For Toronto, the quintiles of NO₂ (ppb) are: ≤-2.9; -2.9-(-1.1); -1.1-0.3; 0.3-2.1; and ≥2.1. For Windsor, the quintiles of NO₂ (ppb) are: ≤1.3; -1.3-(0.4); -0.4-0.5; 0.5-1.3; and ≥1.3.

eTable 4. Estimated average concentration of NO₂ using data from fixed-site monitors during the time period of dense sampling campaigns and across the whole years, by city and year

City	Number of fixed-site stations	Average concentration of NO ₂ at fixed-site stations						
		October 2002	September 2002	May 2004	Estimated average of September 2002 and May 2004	2002	2004	Estimated two-year average of 2002 and 2004
Hamilton	2	20.9	-	-	-	20.9*	17.7*	19.3
Toronto	7	-	22.7	18.7	20.7	20.8	19.5	20.2
Windsor	2	-	-	-	-	19.3	18.1	18.7

* Only one fixed-site station was included because substantial proportion of data ($\geq 30\%$) was missing at a second fixed-site station.

eTable 5. Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and the estimated concentrations of NO₂ from land use regression models (LUR) as well as proximity to major roads and highways, The Ontario Tax Cohort Study, 1982-2004 ^a

Cause of Death	Hamilton RR (95% CI)	Toronto RR (95% CI)	Windsor RR (95% CI)	Pooled estimate RR (95% CI)
<i>All Cardiovascular Disease</i>				
Model 1: NO ₂ from LUR (per 5ppb)	1.12 (1.06 - 1.19)	1.05 (1.00 - 1.09)	1.10 (1.02 - 1.19)	1.08 (1.05 - 1.11)
Model 2: Proximity to roadways ^b	1.02 (0.96 - 1.09)	1.03 (0.95 - 1.12)	1.06 (0.99 - 1.13)	1.04 (1.00 - 1.08)
<i>Ischemic Heart Disease</i>				
Model 1: NO ₂ from LUR (per 5ppb)	1.12 (1.02 - 1.21)	1.06 (1.00 - 1.13)	1.11 (1.00 - 1.23)	1.09 (1.04 - 1.14)
Model 2: Proximity to roadways ^b	1.06 (0.98 - 1.16)	1.12 (1.00 - 1.25)	1.05 (0.97 - 1.15)	1.07 (1.01 - 1.13)
<i>Cerebrovascular Disease</i>				
Model 1: NO ₂ from LUR (per 5ppb)	1.06 (0.92 - 1.22)	0.91 (0.83 - 1.00)	0.96 (0.82 - 1.18)	0.96 (0.90 - 1.05)
Model 2: Proximity to roadways ^b	1.04 (0.90 - 1.20)	0.95 (0.80 - 1.14)	1.02 (0.88 - 1.18)	1.01 (0.92 - 1.10)

^a. Adjusted for age, sex, marriage status (four categories), annual household income (quintiles), and four ecological variables: % of immigrants (quintiles); % of population with less than high school education (continuous); unemployment rate (continuous); and average household income (quintiles).

^b. Exposed: <50m from a major road or <100m from a highway based on subjects' postal code addresses at the time of entry.

eTable 6. Rate ratios (RR) and associated 95% confidence intervals (95% CI) for the association between cause-specific cardiovascular mortality and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, among study participants in Toronto, The Ontario Tax File Cohort Study, 1982-2004. The rate ratios are expressed for an increased of the interquartile range (IQR) of NO₂.^a

Exposure metrics	Year	IQR (ppb)	All cardiovascular disease	Ischemic heart disease	Cerebrovascular disease
			Adjusted RR (95% CI) ^b	Adjusted RR (95% CI)	Adjusted RR (95% CI)
LURs (2002, 2004)	Mean of 2002-2004	4.09	1.04 (1.00 - 1.08)	1.05 (1.00 - 1.11)	0.93 (0.86 - 1.00)
IDW-based extrapolation	1982	4.82	1.02 (0.98 - 1.06)	1.03 (0.98 - 1.10)	0.93 (0.86 - 1.01)
	1992	5.00	1.02 (0.98 - 1.06)	1.05 (0.99 - 1.11)	0.91 (0.84 - 0.99)
	Mean of 1982-1992	4.77	1.02 (0.98 - 1.06)	1.04 (0.99 - 1.10)	0.92 (0.85 - 1.00)
LUR-based extrapolation	1982	5.51	1.03 (1.00 - 1.07)	1.06 (1.00 - 1.12)	0.95 (0.88 - 1.03)
	1992	4.68	1.04 (1.00 - 1.09)	1.06 (1.00 - 1.13)	0.95 (0.87 - 1.04)
	Mean of 1982-1992	4.67	1.04 (1.01 - 1.09)	1.07 (1.01 - 1.13)	0.94 (0.86 - 1.02)
	Mean of 1982-2004	4.40	1.04 (1.00 - 1.08)	1.06 (1.00 - 1.12)	0.94 (0.86 - 1.01)

^a. The RRs for an increase equal to the interquartile range of each predicted distribution of the different exposure metrics to account for differences in the absolute value of the distributions.

^b. The baseline hazard function in the Cox regression models was stratified by 1-year age categories. The model was adjusted for age, sex, marital status (four categories), annual household income (quintiles), and four ecological variables: % of immigrants (quintiles); % of population with less than high school education (continuous); unemployment rate (continuous); and average household income (quintiles).

eTable 7. Rate ratios (RR_{5ppb}) for an increase of 5 ppb in NO₂ and associated 95% confidence intervals (CI) in Toronto for the association between mortality from cardiovascular disease and estimates of nitrogen dioxide evaluated using land use regression models developed for the years of 2002 and 2004 and two back-extrapolation methods for the years of 1982 and 1992, The Ontario Tax Cohort Study, 1982-2004.

Exposure metrics	Year	All cardiovascular disease	Ischemic heart disease	Cerebrovascular disease
		Fully adjusted RR _{5ppb} (95% CI) ^a	Fully adjusted RR _{5ppb} (95% CI)	Fully adjusted RR _{5ppb} (95% CI)
Land-use regression models (2002, 2004)	Mean of 2002-2004	1.05 (1.00 - 1.09)	1.06 (1.00 - 1.13)	0.91 (0.83 - 1.00)
Inverse-distance weighting-based extrapolation	1982	1.02 (0.98 - 1.06)	1.04 (0.98 - 1.10)	0.93 (0.85 - 1.01)
	1992	1.02 (0.98 - 1.06)	1.05 (0.99 - 1.11)	0.91 (0.84 - 0.99)
	Mean of 1982-1992	1.02 (0.98 - 1.06)	1.04 (0.99 - 1.10)	0.92 (0.84 - 1.00)
Land-use regression-based extrapolation	1982	1.03 (1.00 - 1.07)	1.05 (1.00 - 1.10)	0.95 (0.89 - 1.03)
	1992	1.05 (1.00 - 1.09)	1.06 (1.00 - 1.14)	0.95 (0.86 - 1.04)
	Mean of 1982-1992	1.05 (1.01 - 1.09)	1.07 (1.01 - 1.14)	0.94 (0.85 - 1.03)
	Mean of 1982-2004	1.05 (1.01 - 1.09)	1.07 (1.01 - 1.14)	0.93 (0.84 - 1.02)

^a. The baseline hazard function in the Cox regression models was stratified by 1-year age categories. The model was adjusted for age, sex, marriage status (four categories), annual household income (quintiles), and four ecological variables: % of immigrants (quintiles); % of population with less than high school education (continuous); unemployment rate (continuous); and average household income (quintiles).

eTable 8. Correlations between estimated annual mean concentrations of NO₂ (ppb) across three time periods in Toronto, according to the surface maps of NO₂ produced using the land use regression model for 2002-2004 and the two extrapolation methods (based on 5,000 random locations in Toronto), The Ontario Tax File Cohort Study, 1982-2004

	Pearson correlation coefficients between periods ^a		
	1982	1992	2002-2004
LUR ₂₀₀₂₋₂₀₀₄ × ratio of IDW _{fixed} ^b			
1982	1	0.93	0.89
1992		1	0.93
LUR ₂₀₀₂₋₂₀₀₄ × ratio of LUR _{fixed} to LUR _{predicted} ^b			
1982	1	0.78	0.84
1992		1	0.87

^a. The linear relationship between concentrations of NO₂ in two separate years was confirmed from visual inspection of the scatter plots

^b. LUR₂₀₀₂₋₂₀₀₄ is the average of LUR 2002 (R²=0.69) and LUR 2004 (R²=0.71); IDW, inverse distance weighted interpolation; LUR, land use regression model

Reference List

- (1) Statistics Canada. Vital Statistics - Death Database. <http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3233&lang=en&db=imdb&adm=8&dis=2>. Accessed January 20, 2011.
- (2) Goldberg MS, Carpenter M, Theriault G, et al. The accuracy of ascertaining vital status in a historical cohort study of synthetic textiles workers using computerized record linkage to the Canadian Mortality Data Base. *Can J Public Health*. 1993;84:201-204.
- (3) Guibert R. Could the coronary heart disease mortality rates decline be artefactual? *Can J Public Health*. 1991;82:43-45.
- (4) Guibert RL, Wigle DT, Williams JI. Decline of acute myocardial infarction death rates not due to cause of death coding. *Can J Public Health*. 1989;80:418-422.
- (5) Coady SA, Sorlie PD, Cooper LS, et al. Validation of death certificate diagnosis for coronary heart disease: the Atherosclerosis Risk in Communities (ARIC) Study. *J Clin Epidemiol*. 2001;54:40-50.
- (6) Engel LW, Strauchen JA, Chiazze L, Jr., et al. Accuracy of death certification in an autopsied population with specific attention to malignant neoplasms and vascular diseases. *Am J Epidemiol*. 1980;111:99-112.

- (7) Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. *N Engl J Med*. 1985;313:1263-1269.

- (8) Lloyd-Jones DM, Martin DO, Larson MG, et al. Accuracy of death certificates for coding coronary heart disease as the cause of death. *Ann Intern Med*. 1998;129:1020-1026.

- (9) Hoffmann B, Moebus S, Kroger K, et al. Residential exposure to urban air pollution, ankle-brachial index, and peripheral arterial disease. *Epidemiology*. 2009;20:280-288.

- (10) Baccarelli A, Martinelli I, Pegoraro V, et al. Living near major traffic roads and risk of deep vein thrombosis. *Circulation*. 2009;119:3118-3124.

- (11) Finkelstein MM, Jerrett M, Sears MR. Environmental inequality and circulatory disease mortality gradients. *J Epidemiol Community Health*. 2005;59:481-487.

- (12) Gan WQ, Tamburic L, Davies HW, et al. Changes in Residential Proximity to Road Traffic and the Risk of Death From Coronary Heart Disease. *Epidemiology*. 2010;21:642-649.

- (13) Gehring U, Heinrich J, Kramer U, et al. Long-term exposure to ambient air pollution and cardiopulmonary mortality in women. *Epidemiology*. 2006;17:545-551.
- (14) Hoek G, Brunekreef B, Goldbohm S, et al. Association between mortality and indicators of traffic-related air pollution in the Netherlands: A cohort study. *Lancet*. 2002;360:1203-1209.
- (15) Jerrett M, Finkelstein MM, Brook JR, et al. A cohort study of traffic-related air pollution and mortality in Toronto, Ontario, Canada. *Environ Health Perspect*. 2009;117:772-777.
- (16) Tonne C, Melly S, Mittleman M, et al. A case-control analysis of exposure to traffic and acute myocardial infarction. *Environ Health Perspect*. 2007;115:53-57.
- (17) Chen H, Goldberg MS, Crouse DL, et al. Back-extrapolation of estimates of exposure from current land-use regression models. *Atmos Environ*. 2010;44:4346-4354.
- (18) Su JG, Jerrett M, Beckerman B, et al. Predicting traffic-related air pollution in Los Angeles using a distance decay regression selection strategy. *Environ Res*. 2009;109:657-670.

- (19) Crouse DL, Goldberg MS, Ross NA, et al. Postmenopausal Breast Cancer Is Associated with Exposure to Traffic-Related Air Pollution in Montreal, Canada: A Case-Control Study. *Environ Health Perspect.* 2010;118:1578-1583.
- (20) Jerrett M, Burnett RT, Pope CA 3rd, et al. Spatiotemporal Analysis of Air Pollution and Mortality in California Based on the American Cancer Society Cohort: Final Report. 2011. <http://www.arb.ca.gov/research/rsc/10-28-11/item1dfr06-332.pdf>. Accessed March 5, 2012.
- (21) Kramer U, Herder C, Sugiri D, et al. Traffic-related air pollution and incident type 2 diabetes: results from the SALIA Cohort Study. *Environ Health Perspect.* 2010;118:1273-1279.
- (22) Eeftens M, Beelen R, Fischer P, et al.. Stability of measured and modelled spatial contrasts in NO₂ over time. *Occup Environ Med.* 2011;68:765-770.
- (23) Axelson O. Aspects on confounding in occupational health epidemiology. *Scand J Work Environ Health.* 1978;4:98-102.
- (24) Cornfield J, Haenszel W, Hammond EC, et al. Smoking and lung cancer: recent evidence and a discussion of some questions. *J Natl Cancer Inst.* 1959;22:173-203.

- (25) Steenland K, Greenland S. Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. *Am J Epidemiol.* 2004;160:384-392.
- (26) Villeneuve PJ, Goldberg MS, Burnett RT, et al. Associations between cigarette smoking, obesity, sociodemographic characteristics and remote-sensing-derived estimates of ambient PM_{2.5}: results from a Canadian population-based survey. *Occup Env Med.* 2011;68:920-927.
- (27) Statistics Canada. Canadian Community Health Survey (CCHS). <http://www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3226&lang=en&db=imdb&adm=8&dis=2>. Accessed January 28, 2011.
- (28) Thun MJ, Apicella LF, Henley SJ. Smoking vs other risk factors as the cause of smoking-attributable deaths: confounding in the courtroom. *JAMA.* 2000;284:706-712.
- (29) Sahsuvargolu T, Arain A, Kanaroglou P, et al. A land use regression model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Ontario, Canada. *J Air Waste Manag Assoc.* 2006;56:1059-1069.

- (30) Wheeler AJ, Smith-Doiron M, Xu X, et al. Intra-urban variability of air pollution in Windsor, Ontario--measurement and modeling for human exposure assessment. *Environ Res.* 2008;106:7-16.
- (31) Health Canada. Canadian Guidelines for Body Weight Classification in Adults. 2003. http://www.hc-sc.gc.ca/fn-an/nutrition/weights-poids/guide-ld-adult/bmi_chart_java-graph_imc_java-eng.php. Accessed August 17, 2011.