

Control of Weekly Time Trend in Time-Stratified Case-Crossover Design

Xiaoming Wang¹, Sukun Wang² & Warren Kindzierski³

¹ Health Services Statistical and Analytic Methods, Alberta Health and Services, Edmonton, Canada

² Financial Crime Unit, Bank of Montreal, Ontario, Toronto, Canada

³ School of Public Health, University of Alberta, Edmonton, Canada

Correspondence: Xiaoming Wang, #7235, 2nd Floor, West Wing, Aberhart Centre, 11402 University Avenue, Edmonton, AB., T6G 2J3, Canada

Received: July 18, 2021 Accepted: August 17, 2021 Online Published: August 26, 2021

doi:10.5539/ijsp.v10n5p85

URL: <https://doi.org/10.5539/ijsp.v10n5p85>

Abstract

Case-crossover designs have become widespread in biomedical investigations of transient associations. However, the most popular reference-selection strategy - the time-stratified scheme - may not be an optimum solution to control systematic bias in case-crossover studies. To prove this, we conducted a time series decomposition for daily ozone records and examined the capability of the time-stratified scheme to control for yearly, monthly, and weekly time trends; and observed its failure on the control for the weekly time trend. To solve this issue, we proposed a new logistic regression approach in which we suggest the adjustment for the weekly time trend. We compared the performance of the proposed with that of the traditional method by simulation. We further conducted an empirical study to explore the performance of the new logistic regression approach in examining potential associations between ambient air pollutants and acute myocardial infarction hospitalizations. The time-stratified scheme provides effective control for yearly and monthly time trends but not of the weekly time trend. Uncontrolled weekly time trends could be the dominant source of systematic bias in time-stratified case-crossover studies. In contrast, the proposed logistic regression approach can effectively minimize systematic bias in a case-crossover study.

Keywords: case-crossover design, overlap bias, time-stratified scheme, permutation, bias elimination

1. Introduction

Since the initial work by Maclure (Maclure, 1991), the use of case-crossover designs has become widespread in epidemiological and medical investigations of transient associations between risk factors and adverse health events (Janes, Sheppard & Lumley, 2005a; Delaney & Suissa, 2009; Wang, Wang & Kindzierski, 2019), notably in the area of ambient air pollution research (Yusuf, Hawken, Ôunpun, et al., 2004; Carracedo-Martínez, Taracido, Tobias, et al., 2010; Malig, Green, Basu, & Broadwin, 2013; Wang, Kindzieski & Kaul, 2015a & 2015b; Szyszkowicz & de Angelis, 2021). With a case-crossover design, investigators sample only cases and compare each individual's exposure during a short period (hazard period) just before the onset of a case event with exposure at other times (reference periods) in a referent window (Lu, Symons, Geyh & Zeger, 2008).

Compare to general case-control and cohort study designs, these self-matching designs have an obvious advantage of controlling for all time-invariant confounders, either measured or unmeasured. The challenge of a case-only design comes from how to control for time-varying factors. Like a cohort or a case-control study, confounding in a self-matching design arises when there is an unbalanced matching in determinants between hazard and reference periods, leading to various sources of systematic bias. These sources were previously reviewed by Mittleman and Mostofsky (Mittleman & Mostofsky, 2014).

To control for time-varying factors, various reference-selection strategies have been proposed in the literature. Janes, et al. (Janes, Sheppard & Lumley, 2005a) provided a comprehensive review of these strategies and assessed potential bias associated with them. These strategies aim to limit the reference-selection window to a short period that restricts time-varying confounders to be nearly constant across reference-selection windows (Lu, Symons, Geyh & Zeger, 2008). Ambidirectional (Bateson & Schwartz, 1999) and time-stratified (Lumley & Levy, 2000) are among the top two reference-selection strategies of case-crossover studies in the literature and the latter is claimed to be unbiased (Janes, Sheppard & Lumley, 2005b) and optimal (Mittleman, 2005), and intended to avoid potential bias resulted from time trends in the exposure series as well as specific time-varying confounders.

As pointed out by Wang, et al. (Wang, Wang & Kindzierski, 2019), there are two types of bias in a case-crossover

study. One is the bias in a point estimate of effect from imperfect control of time-varying factors (the confounding issue). The other is estimation bias of the standard error (SE), the so-called overlap bias (Janes Sheppard & Lumley, 2005b; Austin, Flanders & Rothman, 1989; Robins & Pike, 1990) from correlations among observations. One can reduce bias in point estimate by using a shorter reference-select window, which results in higher overlap bias because a shorter reference-selection window leads to higher correlation among cases and their matched references. In their paper, the authors concluded that the time-stratified case-crossover design is by no means a final solution to the overlap bias issue. As a remedy, a calibration strategy based on permutation (Fisher, 1935; Pitman, 1937) was proposed to overcome the overlap bias issue. Although the authors proved that their strategy can effectively reduce estimation bias. However, its computational burden is high, especially when the sample size is large. In addition, the authors did not explain where the bias roots from.

In this paper, we sought a better solution to potential systematic bias in a case-crossover study. We started with scrutinizing the capability of the time-stratified scheme to control for time-varying factors and found its failure on the control for the weekly time trend. We then suggested an additional adjustment for the weekly time trend in a logistic regression model. Via avoiding a calibration algorithm, the proposed simple modeling strategy can dramatically reduce the computational burden from permutation. We hypothesize that it can be competitive to the calibration method, and it can effectively minimize both overlap bias and bias in point estimate.

2. Method

2.1 Overlap Bias

The term overlap bias was first introduced to case-crossover studies by Lumley and Levy (Lumley & Levy, 2000). They observed that this bias is similar to the friend control bias in matched case-control designs (Austin, Flanders & Rothman, 1989; Robins & Pike, 1990). For example, in the time-stratified design for the simulation (subsection 2.5), a hazard day could be viewed as a random sample from a study period, and 3 to 4 reference days are days in the same year, same month, and matching weekday. The selection of reference days is conditional on hazard days. This reference-selection scheme makes the hazard days (synonymous of cases) and reference days (synonymous of friend or sibling controls) to be selected from different strata within the study period, causing a correlation between exposure levels on hazard and reference periods. Austin, et al. (Austin, Flanders & Rothman, 1989) showed that choosing friends or siblings as matched controls for each case can lead to the so-called overlap bias.

Overlap bias originates from correlation among matched cases and controls and leads to estimation bias of SE, and therefore a biased significance test of the exposure effect. We further investigated where comes from the correlation among observations by using an air pollution time-series dataset.

2.2 Limitation of the Time-Stratified Scheme

We used the same publicly available air pollution dataset used by Wang, et al. (Wang, Wang & Kindzierski, 2019). Briefly, daily records of five criteria air pollutants from urban monitoring stations in Edmonton, Alberta, Canada spanning 10 years (from April 1, 2000 to March 31, 2010) were obtained from National Ambient Pollution Surveillance (NAPS) Database (National Air Pollution Surveillance Network, 2013). The five pollutants were carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO₂), ozone (O₃), and particulate matter with an aerodynamic diameter ≤ 2.5 (PM_{2.5}). Because different pollutants have different concentration units, each of the five pollutant measures was standardized by the interquartile range (IQR, difference between the 75th and the 25th concentration percentile).

To examine the limitations of the time-stratified reference-select scheme, we conducted a time series decomposition for daily records of O₃. From the daily time series of O₃, we calculated its yearly average (O₃^y), monthly average after removing the yearly time trend (O₃^m), weekly average after removing the yearly, and monthly time trends (O₃^w) and daily average after removing the yearly, monthly and weekly time trends (O₃^d). The above process is an usual time series decomposition, such that $O_3 = O_3^y + O_3^m + O_3^w + O_3^d$. We have depicted the time series and its components in Figure 1.

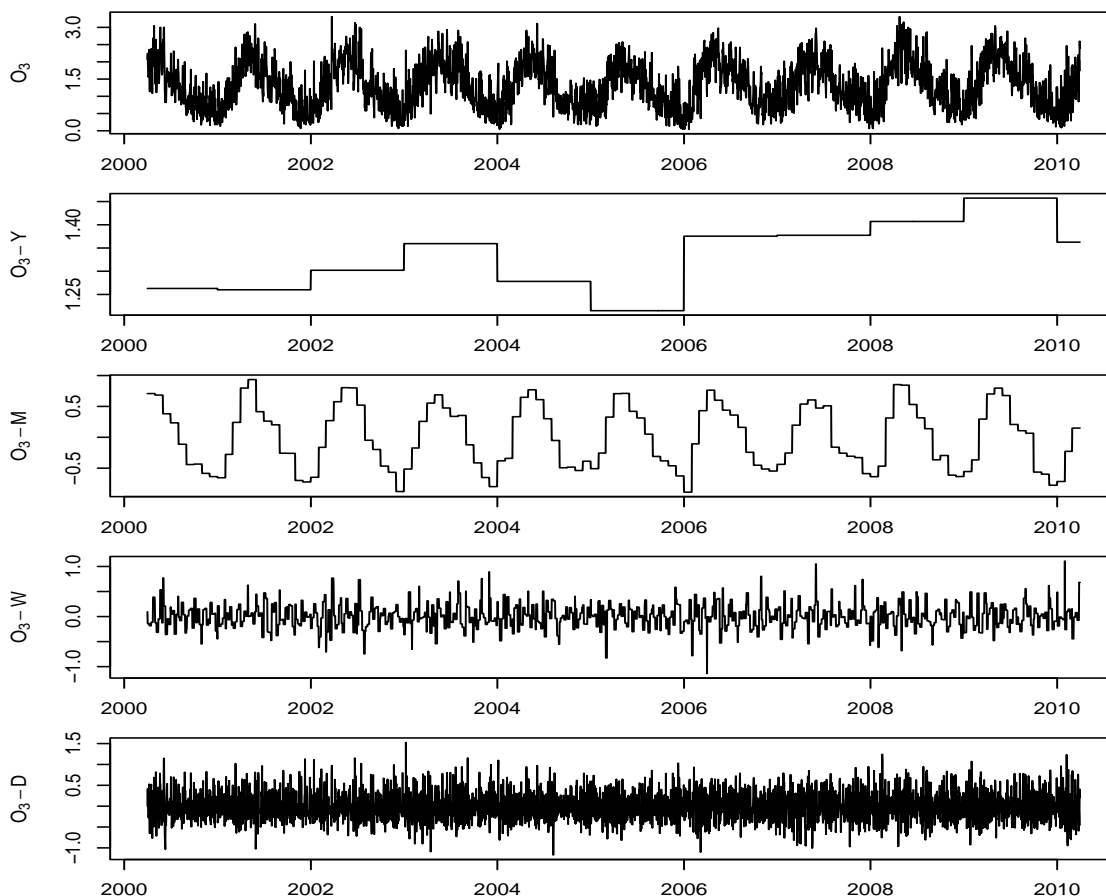


Figure 1. Time series decomposition of daily records of ozone (O_3) from April 1, 2000 to March 31, 2010. Daily records of O_3 (in Edmonton urban area, Canada) are depicted in the top panel. Its yearly, monthly, weekly, and daily components of time trends are depicted in the 2nd to 5th panels, respectively

From Figure 1, we see that: (a) the time-stratified scheme can control for the yearly and monthly time trends of O_3 because a case and its matched references (with 3 to 4 references) are always in the same year and the same month, in which yearly and monthly average values of O_3 are constant; and (b) the time-stratified scheme can do nothing on control of the weekly time trend of O_3 .

The potential transient effect of an air pollutant on a health outcome is a within-day or within-a few-days impact of the air pollutant. It should not include any effect from a weekly time trend. We believe that a lack of control of the weekly time trend is the dominant source of an estimation bias and the overlap bias. Correlations among observations are unignored if the weekly trend time-series is highly correlated.

2.3 Control of the Weekly Time Trend

Suppose that Y is a response variable where $Y = 0/1$ indicates exposure to a referent/hazard period in a case-crossover study. The following traditional logistic regression model (model 1) is the typical analytic tool being used in literature to assess a transient association between adverse events and exposure to a hazard (an air pollutant in our case):

$$\text{logit}(p) = \alpha + \beta O_3, \tag{1}$$

here $\text{logit}(p) = \log(p/(1 - p))$ with $p = Pr(Y = 1)$ and β is the transient effect of O_3 under estimation. Again, we use O_3 as example of the air pollutant. For simplicity we consider a model without covariates.

We proposed the following analytic methods. Given a suitable model adjustment for the weekly trend, we sought a better point estimate and a better significance test. The proposed logistic regression model (model 2) is one adjusted for the weekly time trend of O_3 , i.e.,

$$\text{logit}(p) = \alpha + \beta O_3 + \gamma_1 O_3^y + \gamma_2 O_3^m + \gamma_3 O_3^w. \tag{2}$$

here γ_1 , γ_2 , and γ_3 represent effects of yearly, monthly, and weekly time trends, respectively. Adjustment for yearly and monthly time trends is because some cases will match with three references, and the others will match with four references. It leads to a possible unbalanced matching of yearly and monthly time trends in the hazard and reference periods. Given each of the cases match with the same number of controls in a time-stratified study design, the yearly and monthly terms in model 2 can be removed.

Another logistic regression model (model 3) is designed to directly estimate impact of the daily average (O_3^d) and adjust for the weakly time trend:

$$\text{logit}(p) = \alpha + \beta O_3^d + \gamma_1 O_3^y + \gamma_2 O_3^m + \gamma_3 O_3^w. \tag{3}$$

A transient impact from O_3 is considered to be the effect of O_3^d . It is worth pointing out that models 2 and 3 are equivalent to each other.

Without control of the weekly time trend, correlation among the weekly-average records will enter into the model errors, leading to a relatively higher correlation among the model errors. On the other hand, with control of yearly, monthly, and weekly time trends, correlation among the daily-average records could be quite low for days in the same month but with one week or multiple weeks apart. From this analysis, we conclude that the both the overlap bias and the bias in point estimate root in the same source, i.e., the uncontrolled weekly time trend.

2.4 Calibration Using Permutation

From a statistical point of view, the proposed models 2 and 3 are only for “univariate analysis”. However, bias could also come from other potential time-varying confounders. An ideal strategy is to observe those time-varying confounders and adjust for them. In this way, model 2 can be rewritten as:

$$\text{logit}(p) = \alpha + \beta O_3 + \gamma_1 O_3^y + \gamma_2 O_3^m + \gamma_3 O_3^w + \mu Z. \tag{4}$$

Here $\mu Z = \mu_1 Z^y + \mu_2 Z^m + \mu_3 Z^w$ represents suitable adjustment for time trend from all observed time-varying covariates.

A big challenge comes from unobserved confounders. It is impossible and worthless for investigators to account for all potential time-varying confounders. A possible (but imperfect with computational burden) solution to this challenge is the calibration method suggested by Wang, et al. (Wang, Wang & Kindzierski, 2019). Please referred to the article for details of the calibration algorithm. The basic assumption of the calibration method is that estimated effect ($\hat{\beta}$) can be divided into design bias (b), true effect (β), and random error (e), i.e.

$$\hat{\beta} = \beta + b + e. \tag{5}$$

It implies that regardless of whether the true effect (β) is zero or not, the design bias b will be a constant¹.

Under this assumption, theoretically, we can correctly estimate the design bias b by permutation and completely eliminate it from point estimate $\hat{\beta}$. An unbiased p-value can be obtained by comparing the estimated effect from original data and null effects from permuted samples. Although this assumption may be questionable in practice, we still believe this calibration technique can be beneficial to us for bias reduction in both point estimate and significance test. We will show its performance by simulation.

2.5 Simulation Study Design

We use O_3 as an example to describe our simulation study design. The same design applied to all five air pollutants. Based on the 3,652-day study period (April 1, 2000 to March 31, 2010), we simulated hazard days of adverse events to create different scenarios in order to show the performance of the following modeling strategies.

- Traditional logistic model 1;
- Proposed logistic model 2;
- Traditional model 1 with calibration;
- Proposed model 2 with calibration.

For each scenario, we replicated 5,000 experiments. For each of these experiments: (a) 5,000 hazard days were sampled randomly with replacement from the study period with a probability proportional to $\exp(\beta O_3^d + \gamma_1 O_3^y + \gamma_2 O_3^m + \gamma_3 O_3^w)$; (b)

¹Assumption (5) is the base of unbiasedness of coefficient estimate for the calibration method. The permeation test can be unbiased without the assumption.

reference days for each hazard day were selected using the time-stratified scheme, and (c) data were analyzed using the above four modeling strategies. All analyses were conducted in R (version 4.1.0, 2021-05-18).

We used the null model ($\beta = 0$) to compare the size (under the criterion $\alpha_0 = 0.05$) of the tests for each modeling strategy. All simulation scenarios under the null hypothesis used settings of $\gamma_1 = \gamma_2 = \gamma_3 = \gamma = 0$ or 0.2^2 . The purpose of the simulation analysis was to examine whether type-I errors of the case-crossover model scenarios were close to the nominal level of 50 in 1,000 replications.

We then used the alternative model to compare the power (under the criterion $\alpha_0 = 0.05$) of the tests for each modeling strategy. Scenarios for this check were simulated by setting $\gamma_1 = \gamma_2 = \gamma_3 = \gamma = 0$ and β from 0, 0.1, \dots , 0.3. The purpose of this simulation analysis was to examine rejection rates of significance tests in 5,000 experiments under each scenario.

2.6 Empirical Study Design

Using Alberta Health administrative databases, we obtained historical records of hospital admissions for acute myocardial infarction (AMI, ICD-10 code I21-I22 or ICD-9 code 410) for Edmonton urban dwellers for the 10 years (April 1, 2000 to March 31, 2010). In total 9,811 admissions for AMI occurred for the 10 years among patients aged 20 or over and living in the metropolitan areas of Edmonton.

We explored potential transient associations between AMI hospitalizations and the five air pollutants (the same data used in the simulation study). Cohort and ten sub-cohorts populations were defined using the following admission characteristics:

- the whole cohort (Whole),
- males (Males),
- females (Females),
- subjects aged ≥ 65 (Age ≥ 65),
- subjects aged < 65 (Age < 65),
- subjects categorized as ST-Segment Elevation Myocardial Infarction upon hospitalization (STEMI),
- subjects categorized as Non-ST Segment Elevation Myocardial Infarction upon hospitalization (NSTEMI),
- subjects with pre-existing diabetes (Diabetes),
- subjects with pre-existing hypertension (Hypertension),
- subjects with a pre-existing irregular heartbeat (Dysrhythmia), and
- subjects with the prehistory of heart disease (PIHD).

Records of the whole cohort and each sub-cohort were further divided into temporal subgroups defined by seasons (All season, Spring, Summer, Autumn, and Winter). For each subgroup, (1) hazard days were selected as the k th-day pre-case events ($k = 0 - 4$), and (2) reference days for each hazard day were selected using the time-stratified scheme. These data were analyzed using the traditional logistic regression model 1 and the proposed logistic regression model 2.

3. Results

3.1 Simulated Study Results

Bias check and size check under the null model (with the setting of $\beta = 0$ and $\gamma = 0$ or 0.2) were conducted for each of the five pollutants. Point estimates from each of the four modeling strategies are depicted in Figure 2 and Figure 3, and rejection rates (size, under the criterion $\alpha_0 = 0.05$) of significance tests in 5,000 experiments are reported in Table 1.

²The setting of $\gamma = 0.2$ was based on the empirical analysis, in which we estimated the effects of weekly time trends using model 2 for the 1,375 association analyses. The range of the 1,375 estimated effects is from -0.798 to 0.842.

Table 1. Size of significance tests for model 1-2 (before calibration and after calibrations)

Model	γ	Calibration	CO	NO	NO ₂	O ₃	PM _{2.5}
Model 1	0.0	Before	0.38%	1.04%	1.62%	0.68%	4.08%
Model 1	0.0	After	5.50%	4.78%	5.62%	5.06%	5.54%
Model 2	0.0	Before	4.34%	3.80%	4.08%	5.72%	5.28%
Model 2	0.0	After	5.12%	5.40%	5.54%	5.08%	5.44%
Model 1	0.2	Before	62.28%	97.58%	61.58%	9.18%	99.34%
Model 1	0.2	After	4.96%	5.82%	6.18%	4.66%	5.50%
Model 2	0.2	Before	4.24%	3.70%	4.14%	5.16%	5.16%
Model 2	0.2	After	5.24%	5.58%	5.72%	5.88%	6.10%

Note: Model 1=traditional logistic regression model on pollutant without adjustment; Model 2=proposed logistic regression model on pollutant with adjustment on its yearly, monthly and weekly time trend; each size (under the criterion $\alpha_0 = 0.05$) was calculated from 5000 independent replication of experiments

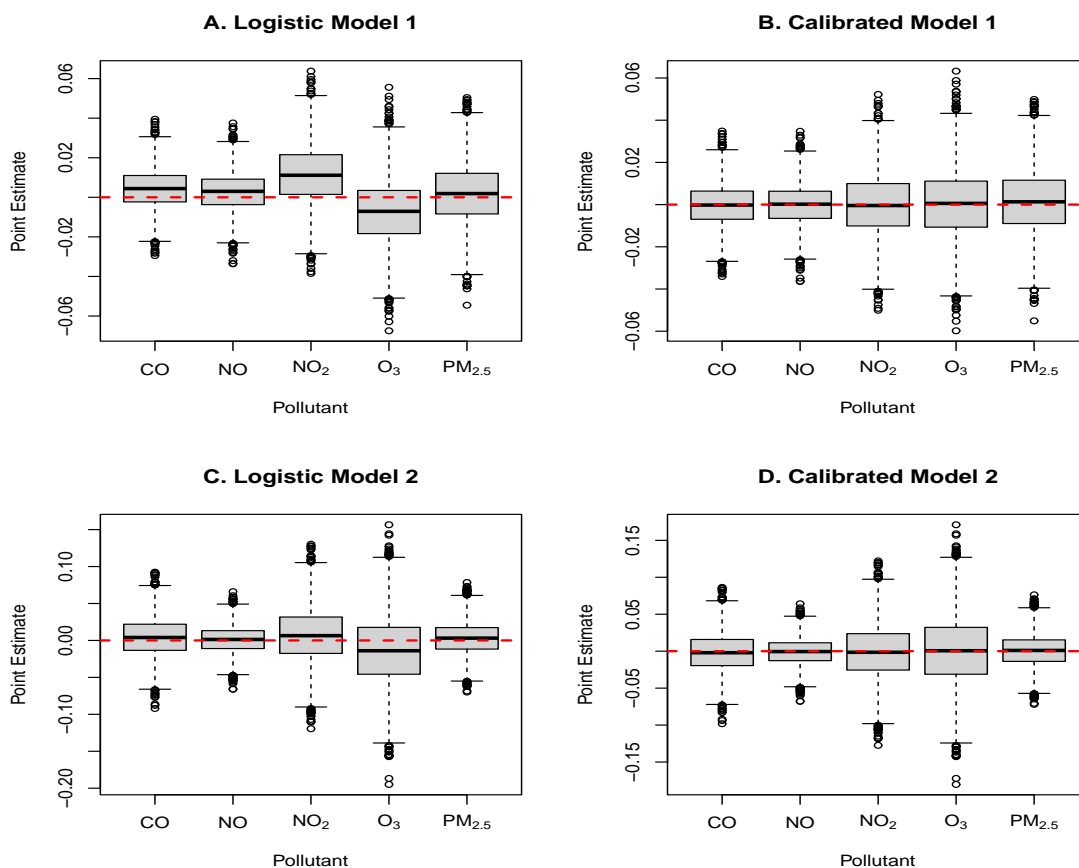


Figure 2. Boxplot of distribution the estimated biases of effect of one of the five pollutants using one of the four analytic modeling strategies respectively. Scenarios are simulated under $\beta = 0$ and $\gamma = 0$, and study period is the whole-cycle-period from April 1, 2000 to March 31, 2010. The five pollutants are carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO₂), ozone (O₃), and particulate matter with an aerodynamic diameter ≤ 2.5 (PM_{2.5})

From the top panel of Table 1, we can see that in the scenarios with $\gamma = 0$, the size of model 1 (before calibration) is far less from the nominal level 5% in 5,000 replications (i.e., the 0.05 criterion), especially for CO, NO, and O₃. Considering the

point estimates in these situations are almost unbiased (Figure 2-panel A), we can conclude that the bias in size is mainly from correlation among model errors (i.e., overlap bias). It implies that, without control of the weekly time trend, model 1 provides overestimates of standard errors, leading to higher overlap bias. On the other hand, with suitable adjustment of the weekly time trend (model 2), the overlap bias is almost eliminated and significance tests are nearly unbiased. It appears that the calibration method provides reasonable and robust significance tests for both models.

The bottom panel of Table 1 shows that under the setting of $\gamma = 0.2$ there is a big problem in using model 1 (before calibration). The large size of significance tests in model 1 (before calibration) is from the uncontrolled weekly time trend, leading to biased point estimates (refer to Figure 3-panel A) as well as a higher correlation among errors of the model. In this scenario, we observe again the superpower of the calibration strategy in bias correction (refer to Figure 3-panel B). The proposed logistic regression model 2 is nearly unbiased in significance tests both before and after calibration. The superior performance of model 2 roots from thorough control of the weekly time trend. In this case, the calibration method can do very little in further reducing bias for model 2.

Figure 2 represents the simulated performance in point estimates of the traditional logistic model 1 vs the proposed logistic model 2 under the setting of $\gamma = 0$. Because of the null hypothesis, a point estimate is its estimation bias. In this scenario point estimates from model 1 are nearly unbiased (Figure 2-panel A). After calibration (Figure 2-panel B), bias in point estimates is virtually eliminated. Model 2 provides very similar results in estimation bias before and after calibration (Figure 2-panel C and D).

Comparing the top panels (A and B) and the bottom panels (C and D) in Figure 2, we can see a difference in the variance of estimation bias: the range of top panels is roughly within ± 0.06 , while that of bottom panels is roughly within ± 0.15 . This can be explained by the difference between models 1 and 2. In model 1, the time-stratified scheme is used to control for yearly and monthly time trends, so the estimated effect is from the weekly time trend (i.e. weekly average after removing yearly and monthly time trends). In contrast, with control of yearly, monthly, and weekly time trends, the estimated effect from model 2 is that of the daily time trend (daily time series after removing yearly, monthly, and weekly time trends). The target estimation of model 1 represents the effect of the weekly time trend effect, instead of the transient effect under investigation.

Model 1 and model 2 performed quite differently in point estimates for $\gamma = 0.2$, (Figure 3). We see that: (a) without control of the weekly time trend, model 1 could be unacceptably biased in the coefficient estimate (Figure 3-panel A); (b) with thorough control of the weekly time trend, model 2 could be almost unbiased in point estimates (refer to Figure 3-panel C); (c) the calibration method can virtually eliminate bias from model 1 by permutation (refer to Figure 3-panel B); and (d) model 2 and the calibrated model 2 perform similarly (refer to Figure 3-panel C and D), suggesting that model 2 is sufficient in point estimate and further calibration is unnecessary in the simulated scenarios.

Bias check and power check under the alternative model (with the setting of $\gamma = 0$ and β from 0, 0.1, \dots , 0.3) were conducted for each of the five pollutants. Rejection rates (power, under the criterion $\alpha_0 = 0.05$) of significance tests in 5,000 experiments under each scenario are reported in Table 2.

We observed the following results from the power check panel of Table 2: (a) the power from model 1 before calibration was the poorest because of the overlap bias; (b) model 2 before calibration provided competitive power results with model 2 with calibration, and (c) model 1 with calibration was better than model 1 before calibration in test power but poorer than model 2 before calibration and model 2 after calibration.

From the bias check panel of Table 2, we see that: (a) model 1 (before or after calibration) gets biased estimates under alternatives; (b) model 2 (before or after calibration) gets nearly unbiased estimation; and (c) in all the experiments, the calibration provides smaller or similar estimation bias, regardless whether the assumption (5) holds or not.

3.2 Empirical Study Results

Estimated effects of 1,375 candidate pollutant-AMI associations (11 cohort/sub-cohorts \times 5 subgroups \times 5 lags \times 5 pollutants) from the traditional logistic model 1 and the proposed logistic model 2 were presented in Supplementary Material files "Est-AMI-application-model1.csv" and "Est-AMI-application-model2.csv", respectively.

In the estimates from model 1, only 793 (582) candidate associations were estimated as positive (negative) effects. Among them, there were 140 (39) candidate associations with p-values ≤ 0.05 (0.01). On the other hand, 634 (741) candidate associations were estimated as positive (negative) effects in the estimates from model 2. Among them there were 75 (10) candidate associations with p-value ≤ 0.05 (0.01). Using Bonferroni correction with a criterion $\alpha_0 = 0.05/1375$, none of the candidate associations from model 1 or model 2 had a p-value less than the Bonferroni correction criterion (0.000036). The correlation between estimates from model 1 and estimates from model 2 was 0.605.

The 39 associations detected by model 1 with p-value ≤ 0.01 are reported in Table 3. From the table we can see an

anomalous seasonal pattern: CO, NO, and NO₂ have positive effects in spring and negative effects in autumn; O₃ has a negative effect in winter and spring, and a positive effect in summer. It does not make sense that a hazard (an air pollutant) has a protective effect in one season while in another season has a harmful effect. This phenomenon suggests that the estimate of the traditional logistic model 1 contains some kind of seasonal trend. We believe this is because that model 1 is unable to control for the weekly time trend of an exposure time series.

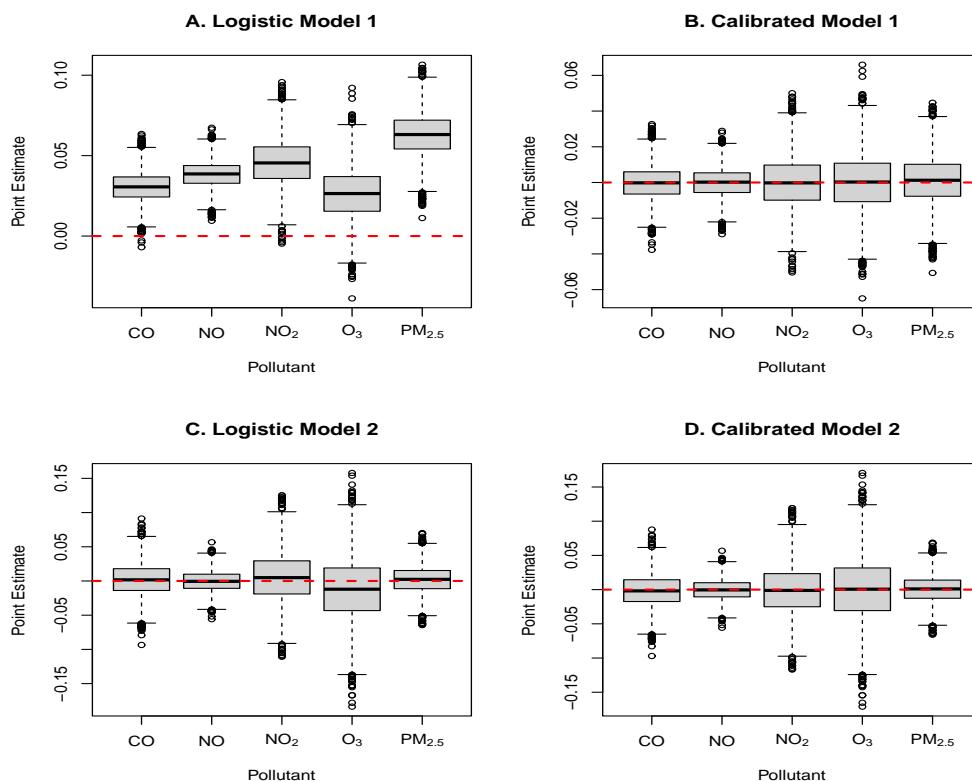


Figure 3. Box plot of distribution of the estimated biases of effect of one of the five pollutants using one of the four analytic modeling strategies respectively. Scenarios are simulated under $\beta = 0$ and $\gamma = 0.2$, and study period is the whole-cycle-period from April 1, 2000 to March 31, 2010. The five pollutants are carbon monoxide (CO), nitrogen monoxide (NO), nitrogen dioxide (NO₂), ozone (O₃), and particulate matter with an aerodynamic diameter ≤ 2.5 (PM_{2.5})

Table 2. Estimation Bias and power of significance test for model 1-2 (before and after calibrations) with $\gamma = 0$

Scenario		Estimation Bias				Power of Test			
pollutant	β	Before Calibration		After Calibration		Before Calibration		After Calibration	
		Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
CO	0.0	0.0043	0.0045	-0.0003	-0.0016	0.38%	4.34%	5.50%	5.12%
CO	0.1	-0.0677	-0.0033	-0.0723	-0.0093	64.70%	96.32%	83.30%	93.80%
CO	0.2	-0.1391	-0.0109	-0.1437	-0.0169	99.98%	100%	100%	100%
CO	0.3	-0.2097	-0.0175	-0.2143	-0.0236	100%	100%	100%	100%
NO	0.0	0.0028	0.0012	-0.0001	-0.0006	1.04%	3.80%	4.78%	5.40%
NO	0.1	-0.0580	-0.0092	-0.0608	-0.0111	98.52%	99.96%	99.08%	99.94%
NO	0.2	-0.1190	-0.0190	-0.1218	-0.0209	100%	100%	100%	100%
NO	0.3	-0.1820	-0.0284	-0.1848	-0.0302	100%	100%	100%	100%
NO ₂	0.0	0.0114	0.0069	-0.0002	-0.0010	1.62%	4.08%	5.62%	5.54%
NO ₂	0.1	-0.0607	-0.0034	-0.0722	-0.0114	41.88%	70.38%	51.76%	67.44%
NO ₂	0.2	-0.1329	-0.0132	-0.1445	-0.0211	96.24%	99.86%	97.94%	99.82%
NO ₂	0.3	-0.2057	-0.0236	-0.2173	-0.0315	100%	100%	100%	100%
O ₃	0.0	-0.0074	-0.0138	0.0003	0.0007	0.68%	5.72%	5.06%	5.08%
O ₃	0.1	-0.0811	-0.0166	-0.0734	-0.0021	3.56%	43.08%	37.86%	55.64%
O ₃	0.2	-0.1544	-0.0190	-0.1467	-0.0044	46.22%	96.92%	90.92%	98.66%
O ₃	0.3	-0.2275	-0.0217	-0.2198	-0.0071	93.98%	100%	99.88%	100%
PM _{2.5}	0.0	0.0019	0.0031	0.0013	0.0010	4.08%	5.28%	5.54%	5.44%
PM _{2.5}	0.1	-0.0411	0.0034	-0.0417	0.0013	98.18%	99.94%	98.06%	99.88%
PM _{2.5}	0.2	-0.0820	0.0021	-0.0826	0.0000	100%	100%	100%	100%
PM _{2.5}	0.3	-0.1206	0.0001	-0.1211	-0.0020	100%	100%	100%	100%

Note: Model 1=traditional logistic regression model on pollutant without adjustment; Model 2=proposed logistic regression model on pollutant with adjustment on its yearly, monthly and weekly time trend; each power (under the criterion $\alpha_0 = 0.05$) was calculated from 5000 independent replication of experiments

Table 3. Detected associations with p-value ≤ 0.01 between pollutants and AMI hospitalization by the traditional logistic model 1

Sub-cohort	Season	N	Pollutant	Lag (day)	Odds Ratio (95% CI)
Whole	Spring	2505	CO	4	1.068 (1.017, 1.121)
Age<65	Spring	1045	CO	4	1.121 (1.042, 1.205)
NSTEMI	Spring	1397	CO	4	1.093 (1.022, 1.168)
Whole	Autumn	2408	CO	3	0.919 (0.865, 0.976)
Age≥65	Autumn	1320	CO	3	0.893 (0.822, 0.971)
Whole	Spring	2505	NO	2	1.067 (1.021, 1.115)
Whole	Spring	2505	NO	3	1.066 (1.020, 1.114)
Whole	Spring	2505	NO	4	1.068 (1.019, 1.119)
Male	Spring	1631	NO	3	1.081 (1.024, 1.141)
Male	Spring	1631	NO	4	1.078 (1.019, 1.14)
Age<65	Spring	1045	NO	3	1.106 (1.034, 1.184)
Age<65	Spring	1045	NO	4	1.110 (1.034, 1.191)
STEMI	Spring	1108	NO	3	1.091 (1.024, 1.163)
Dysrhythmia*	Spring	428	NO	3	1.158 (1.046, 1.282)
NSTEMI	Spring	1397	NO	4	1.091 (1.024, 1.162)
Whole	Autumn	2408	NO	3	0.915 (0.861, 0.973)
Whole	Autumn	2408	NO	4	0.907 (0.852, 0.966)
Male	Autumn	1607	NO	2	0.906 (0.841, 0.976)
Male	Autumn	1607	NO	4	0.899 (0.832, 0.972)
Age≥65	Autumn	1320	NO	3	0.880 (0.806, 0.96)
Whole	Spring	2505	NO ₂	2	1.093 (1.030, 1.160)
STEMI	Spring	1108	NO ₂	2	1.138 (1.04, 1.244)
Whole	Autumn	2408	NO ₂	3	0.837 (0.761, 0.921)
Whole	Autumn	2408	NO ₂	4	0.873 (0.794, 0.960)
Male	Autumn	1607	NO ₂	2	0.857 (0.764, 0.96)
Male	Autumn	1607	NO ₂	3	0.805 (0.716, 0.905)
Male	Autumn	1607	NO ₂	4	0.840 (0.748, 0.945)
Age≥65	Autumn	1320	NO ₂	3	0.828 (0.727, 0.944)
HTN	Autumn	1423	NO ₂	3	0.841 (0.743, 0.953)
NSTEMI	Autumn	1312	NO ₂	3	0.800 (0.703, 0.911)
NSTEMI	Autumn	1312	NO ₂	4	0.830 (0.729, 0.946)
Whole	Winter	2475	O ₃	1	0.870 (0.783, 0.967)
Whole	Spring	2505	O ₃	4	0.882 (0.815, 0.955)
Female	Spring	874	O ₃	4	0.832 (0.726, 0.954)
NSTEMI	Spring	1397	O ₃	4	0.870 (0.783, 0.966)
Whole*	Summer	2423	O ₃	2	1.144 (1.037, 1.261)
NSTEMI*	Summer	1288	O ₃	2	1.204 (1.052, 1.377)
HTN*	Summer	1433	O ₃	2	1.183 (1.041, 1.345)
Age<65	Spring	1045	PM _{2.5}	4	1.107 (1.028, 1.191)

Note: N=number of observations; CI=confidence interval; STEMI = ST Segment Elevation Myocardial Infarction; NSTEMI = Non-ST Segment Elevation Myocardial Infarction; HTN = Hypertension; PIHD = prehistory of heart disease; *Detected by both the traditional logistic model 1 and the proposed logistic model 2.

Table 4. Detected associations with p-value ≤ 0.01 between pollutants and AMI hospitalization by the proposed logistic model 2

Sub-cohort	Season	N	Pollutant	Lag (day)	Odds Ratio (95% CI)
Dysrhythmia	Spring	428	CO	3	1.319 (1.094, 1.591)
Dysrhythmia*	Spring	428	NO	3	1.210 (1.048, 1.397)
HTN	All Season	5848	NO ₂	0	0.912 (0.850, 0.978)
Whole*	Summer	2423	O ₃	2	1.216 (1.061, 1.393)
NSTEMI*	Summer	1288	O ₃	2	1.308 (1.086, 1.576)
Diabetes	All Season	2667	O ₃	3	1.192 (1.047, 1.356)
HTN*	Summer	1433	O ₃	2	1.281 (1.072, 1.531)
PIHD	All Season	3076	O ₃	3	1.174 (1.040, 1.325)
STEMI	Autumn	1096	PM _{2.5}	4	1.156 (1.047, 1.276)
Dysrhythmia	Spring	428	PM _{2.5}	3	1.247 (1.058, 1.468)

Note: N=number of observations; CI=confidence interval; STEMI = ST Segment Elevation Myocardial Infarction; N-STEMI = Non-ST Segment Elevation Myocardial Infarction; HTN = Hypertension; PIHD = prehistory of heart disease; SE=standard error; *Detected by both the traditional logistic model 1 and the proposed logistic model 2.

On the other hand, with adjustment for the weekly time trend, model 2 tells a different story. The 10 associations detected by model 2 with p-value ≤ 0.01 are reported in Table 4. From the table, we do not observe any anomalous seasonal pattern as shown in Table 3. All pollutant-AMI associations were consistent, irrespective of season: nine estimated effects for CO, NO, O₃ or PM_{2.5} were positive, and 1 estimated effect of NO₂ was negative.

4. Discussion

Key points of the paper are summarized as follows:

- Using the ozone time series decomposition, we concluded that the time-stratified reference-select scheme lacks capability on control of weekly time trend;
- New strategies were proposed to solve the bias challenge;
 - Modified logistic regression to control for weekly time trend (recommended);
 - Calibration approach for control of bias from unobserved covariates;
- Comparison was done by simulation and empirical studies between the traditional logistic regression and the proposed approach;
 - The simulation study showed that control of weekly time trend can eliminate potential overlap bias as well as estimation bias;
 - The empirical study showed the new approach provides more reasonable association findings compare to the traditional approach.

Using decomposition of an ozone time series dataset, we scrutinized the capability of the time-stratified reference selection scheme to control for yearly, monthly, and weekly time trends when used in case-crossover studies. This scheme provides effective control of yearly and monthly time trends but not of a weekly time trend. We believe that lack of control of the weekly time trend is the dominant source of systematic bias in the existing case-crossover studies using the time-stratified reference selection scheme.

An inability to control for the weekly time trend will lead to overlap bias and possible bias in a point estimate. Modeling adjustment for the weekly time trend can be an effective way to reduce this systematic bias. Further model adjustments or calibration may be necessary if other potential time-varying confounders (observed or unobserved) have to be taken into account.

Through simulation, we show that the typical case-crossover study design (i.e., a time-stratified reference selection scheme combined with a traditional logistic regression) leads to an estimate of the weekly time trend effect, rather than the transient effect under investigation. The difference in results could be big between the typical approach and the proposed method. Based on it, we recommend that researchers be cautious and make necessary adjustments for the weekly time trend in both exposures and covariates in their case-crossover application studies.

Acknowledgements

Ethical approval for the study was granted by a University Health Research Ethics Board-Health Panel. Patient records/information was anonymized and de-identified with a unique scrambled ID and released to us in this form by the ministry of health prior to analysis.

Conflicts of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

- Austin, H., Flanders, W. D., & Rothman, K. J. (1989). Bias arising in case-control studies from selection of controls from overlapping groups. *International Journal of Epidemiology*, *18*, 713-716. <https://doi.org/10.1093/ije/18.3.713>
- Bateson, T. F., & Schwartz, J. (1999). Control for seasonal variation and time trend in case crossover studies of acute effects of environmental exposures. *Epidemiology*, *10*, 539-544.
- Carracedo-Martínez, E., Taracido, M., Tobias, A., Saez, M., & Figueiras, A. (2010). Case-Crossover Analysis of Air Pollution Health Effects: A Systematic Review of Methodology and Application. *Environmental Health Perspectives*, *118*(8), 1173-1182. <https://doi.org/10.1289/ehp.0901485>
- Delaney, J. A., & Suissa, S. (2009). The case-crossover study design in pharmacoepidemiology. *Statistical Methods in Medical Research*, *18*(1), 53-65. <https://doi.org/10.1177/0962280208092346>
- Fisher, R. A. (1935). *The design of experiments*. New York: Hafner.
- Janes, H., Sheppard, L., & Lumley, T. (2005a). Case-crossover analyses of air pollution exposure data: referent selection strategies and their implications for bias. *Epidemiology*, *16*(6), 717-726. <https://doi.org/10.1097/01.ede.0000181315.18836.9d>
- Janes, H., Sheppard, L., & Lumley, T. (2005b). Overlap bias in the case-crossover design, with application to air pollution exposures. *Statistics in Medicine*, *24*(2), 285-300. <https://doi.org/10.1002/sim.1889>
- Lu, Y., Symons, J. M., Geyh, A. S., & Zeger, S. L. (2008). An approach to checking case-crossover analyses based on equivalence with time-series methods. *Epidemiology*, *19*(2), 169-175. <https://doi.org/10.1097/EDE.0b013e3181632c24>
- Lumley, T., & Levy, D. (2000). Bias in the case-crossover design: implications for studies of air pollution. *Environmetrics*, *11*, 689-704. [https://doi.org/10.1002/1099-095X\(200011/12\)11:6h689::AID-ENV439i3.0.CO;2-N](https://doi.org/10.1002/1099-095X(200011/12)11:6h689::AID-ENV439i3.0.CO;2-N)
- Maclure, M. (1991). The case-crossover design: a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology*, *133*(2), 144-153. <https://doi.org/10.1093/oxfordjournals.aje.a115853>
- Malig, B. J., Green, S., Basu, R., & Broadwin, R. (2013). Coarse Particles and Respiratory Emergency Department Visits in California. *American Journal of Epidemiology*, *178*(1), 58-69. <https://doi.org/10.1093/aje/kws451>
- Mittleman, M. A. (2005). Optimal Referent Selection Strategies in Case-Crossover Studies: A Settled Issue. *Epidemiology*, *16*, 715-716. <https://doi.org/10.1097/01.ede.0000183170.92955.25>
- Mittleman, M. A., & Mostofsky, E. (2014). Exchangeability in the case-crossover design. *International Journal of Epidemiology*, *43*(5), 1645-1655. <https://doi.org/10.1093/ije/dyu081>
- National Air Pollution Surveillance Network. (2013). Environment Canada. <http://www.etc-cte.ec.gc.ca/NapsData/Default.aspx>
- Pitman, E. J. G. (1937). Significance tests which may be applied to samples from any population. *Royal Statistical Society Supplement*, *4*, 119-130 and 225-232 (parts I and II).
- Robins, J., & Pike, M. (1990). The validity of case-control studies with nonrandom selection of controls. *Epidemiology*, *1*(4), 273-284. <https://doi.org/10.1097/00001648-199007000-00004>
- Szyszkowicz, M., & de Angelis, N. (2021). Ambient air pollution and emergency department visits in Toronto, Canada. *Environmental Science and Pollution Research*, *28*, 28789C28796(2021). <https://doi.org/10.1007/s11356-021-12519-3>
- Wang, X., Kindziarski, W., & Kaul, P. (2015). Comparison of transient associations of air pollution and AMI hospitalization in two cities of Alberta, Canada, using a case-crossover design. *BMJ Open*, *5*(11), e009169.

<https://doi.org/10.1136/bmjopen-2015-009169>

Wang, X., Kindzierski, W., & Kaul, P. (2015). Air Pollution and Acute Myocardial Infarction Hospital Admission in Alberta, Canada: A Three-Step Procedure Case-Crossover Study. *PLoS One*, *10*(7), e0132769.

<https://doi.org/10.1371/journal.pone.0132769>

Wang, X., Wang, S., & Kindzierski, W. (2019). Eliminating Systematic Bias from Case-Crossover Designs. *Statistical Methods in Medical Research*, *28*(10-11), 3100-3111.

<https://doi.org/10.1177/0962280218797145>

Yusuf, S., Hawken, S., Ôunpuu, S., et al. (2004). Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet*, *364*(9438), 937-952.

[https://doi.org/10.1016/S0140-6736\(04\)17018-9](https://doi.org/10.1016/S0140-6736(04)17018-9)

Appendix A

Est-AMI-application-model1.csv

Estimated effects of 1,375 candidate pollutant-AMI associations from the traditional logistic model 1 in the empirical study were presented in file "Est-AMI-application-model1.csv".

Appendix B

Est-AMI-application-model2.csv

Estimated effects of 1,375 candidate pollutant-AMI associations from the proposed logistic model 2 in the empirical study were presented in file "Est-AMI-application-model2.csv".

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).