

UC San Diego

UC San Diego Previously Published Works

Title

Differential Participation, a Potential Cause of Spurious Associations in Observational Cohorts in Environmental Epidemiology.

Permalink

<https://escholarship.org/uc/item/2v86378j>

Journal

Epidemiology, 35(2)

Authors

Chen, Chen

Chen, Hong

Kaufman, Jay

et al.

Publication Date

2024-03-01

DOI

10.1097/EDE.0000000000001711

Peer reviewed

Differential Participation, a Potential Cause of Spurious Associations in Observational Cohorts in Environmental Epidemiology

 Chen Chen,^a Hong Chen,^{b,c,d,e,*} Jay S. Kaufman,^f and Tarik Benmarhnia^{a,*}

Abstract: Differential participation in observational cohorts may lead to biased or even reversed estimates. In this article, we describe the potential for differential participation in cohorts studying the etiologic effects of long-term environmental exposures. Such cohorts are prone to differential participation because only those who survived until the start of follow-up and were healthy enough before enrollment will participate, and many environmental exposures are prevalent in the target population and connected to participation via factors such as geography or frailty. The relatively modest effect sizes of most environmental exposures also make any bias induced by differential participation particularly important to understand and account for. We discuss key points to consider for evaluating differential participation and use causal graphs to describe two example mechanisms through which differential participation can occur

in health studies of long-term environmental exposures. We use a real-life example, the Canadian Community Health Survey cohort, to illustrate the non-negligible bias due to differential participation. We also demonstrate that implementing a simple washout period may reduce the bias and recover more valid results if the effect of interest is constant over time. Furthermore, we implement simulation scenarios to confirm the plausibility of the two mechanisms causing bias and the utility of the washout method. Since the existence of differential participation can be difficult to diagnose with traditional analytical approaches that calculate a summary effect estimate, we encourage researchers to systematically investigate the presence of time-varying effect estimates and potential spurious patterns (especially in initial periods in the setting of differential participation).


Keywords: Adjusted survival curve; Differential participation; Environmental epidemiology; Observational study; Selection bias

Submitted February 14, 2023; accepted December 18, 2023

From the ^aScripps Institution of Oceanography, University of California San Diego, La Jolla, CA; ^bEnvironmental Health Science and Research Bureau, Health Canada, Ottawa, Ontario, Canada; ^cPublic Health Ontario, Toronto, Ontario, Canada; ^dICES, Toronto, Ontario, Canada; ^eDalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; and ^fDepartment of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada.*Hong Chen and Tarik Benmarhnia are considered as co-senior authors.

Funding for this study was provided by Health Canada (#810630) under the Addressing Air Pollution Horizontal Initiative (AAPHI) of the Government of Canada and National Institute on Aging (RF1AG080948). The authors report no conflicts of interest.

All codes for simulation and relevant analyses can be found at https://github.com/suthlam/differential_participation_simulation.git. All codes for g-computation of the real-life example (Canadian Community Health Survey cohort) can be found at https://github.com/suthlam/cchs_g_computation.git. Statistics Canada's policy on data privacy and confidentiality prohibits the analytical cohort used to be freely available in the article or in a public repository. However, access can be granted through Statistics Canada's Research Data Centre program (<https://www.statcan.gc.ca/en/microdata/data-centres>). Environmental exposures are available upon request to the original authors of the data.

 Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Chen Chen, Scripps Institution of Oceanography, University of California San Diego, 8885 Biological Grade, La Jolla, CA 92037. E-mail: chc048@ucsd.edu.

Copyright © 2023 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

ISSN: 1044-3983/24/352-174184

DOI: 10.1097/EDE.0000000000001711

(*Epidemiology* 2024;35: 174–184)

Most epidemiologic evidence regarding the health impacts of long-term exposures to air pollution—or other environmental exposures—has been obtained through observational cohort studies. However, spurious associations, sometimes counterintuitive and labeled as paradoxes, may commonly occur in such observational settings and be underreported in the published literature. These spurious associations can be due to a discrepancy between the study population and the target population.^{1(chap14)} This lack of representativeness or selection bias has been dismissed by some authors,² but it has the potential to generate misleading associations.³ Differential participation, which occurs frequently in observational cohorts, especially those studying the etiologic effects of long-term environmental exposures, is one likely cause of the discrepancy between the study and target populations.

Differential participation arises from non-participation when subjects in the target population (1) opt not to participate (i.e., nonresponse) or (2) have the outcome of interest before the start of follow-up and thus are not eligible, leading to these outcomes being excluded from the cohort (i.e., left truncation).^{4,5} Since most cohorts only enroll subjects free of the outcome of interest (by design) and healthy enough to participate (in practice), the outcome is generally associated

with the probability of participation directly or via shared risk factors. If the magnitude of nonparticipation does not differ by the exposure of interest (i.e., nondifferential participation), analytical methods such as Cox proportional hazards models can account for it by only including person-times observed in the analysis and will estimate unbiased associations.⁶ However, in real-life settings, participation is unlikely to be independent of long-term environmental exposures, and this can lead to misleading associations due to conditioning on a common effect of exposure and outcome (i.e., collider stratification bias with participation as the collider).⁷

Selection bias caused by differential participation has been described in observational cohort studies of nonenvironmental exposures and is sometimes referred to as left truncation,^{8,9} left informative censoring,^{10,11} non-representative,¹² nonresponse bias,^{5,13,14} healthy worker bias,⁴ survival bias,¹⁵ live-birth bias,¹⁶ index event bias,¹⁷ or simply as collider bias and selection bias.^{18–21} A well-known example is the paradox in which smoking (i.e., exposure) appears to confer a counterintuitive protective effect against preeclampsia in a pregnancy cohort.^{22,23} This paradox can be explained by collider bias due to differential participation in the target population of all pregnancies—only subjects not experiencing early pregnancy loss were enrolled into the study population and followed for preeclampsia, while early pregnancy loss (i.e., a collider) is both affected by smoking and abnormal placentation, a shared risk factor with preeclampsia. In other words, among cohort participants or pregnant women not experiencing early pregnancy loss, smokers were less likely to have abnormal placentation and subsequent preeclampsia, which leads to a counterintuitive protective effect of smoking on preeclampsia.

Cohorts studying etiologic effects of long-term environmental exposures are especially susceptible to experiencing counterintuitive associations caused by differential participation because environmental exposures are generally prevalent in the target population long before the cohort enrollment and could impact participation via frailty or related geographic factors. Moreover, the expected effects of most environmental exposures are relatively small, making it easier for a reversed association to occur even in the presence of a small bias. However, the literature on differential participation in observational cohorts of environmental epidemiology is limited. Existing publications have mostly focused on specific subpopulations such as birth cohorts and occupational cohorts, in which the bias is caused by a necessary restriction of analysis to live births,^{16,24} pregnancies without early loss,^{22,25} or existing employees.⁴ Yet, using cohorts of the general population may similarly yield counterintuitive associations due to differential participation, but such issues have not been discussed to date.

In this article, we provide an in-depth discussion of differential participation in observational cohorts of the general population and how it might cause spurious or even counterintuitive associations in environmental health research. We also discuss the washout method as one potential analytical

solution to account for such bias without the need for additional data on nonparticipants or shared risk factors between participation and outcome. Specifically, we first discuss key points to consider in evaluating and accounting for differential participation. We also use causal graphs to describe two possible bias mechanisms caused by differential participation (geographic factor-driven and frailty-driven). Next, we demonstrate the existence of a counterintuitive association in a survey-based national observational cohort studying fine particulate matter (PM_{2.5}) and mortality and apply the washout method in this real-life example. Last, we conduct simple simulations to mimic the counterintuitive associations observed in the real-life example and demonstrate the efficacy of the washout analysis.

METHODS

How to Evaluate and Account for Differential Participation

We include a list of key points to consider in evaluating and accounting for differential participation in Table 1. To evaluate whether selection bias exists in an observational cohort, we need to specify the target population, which should be based on the research question of interest (Question 1 of Table 1). In this article, we assume that the target population is a population free of the outcome at cohort initiation. Second, we need to assess whether differential participation may exist between the study cohort and the target population based on our substantive knowledge of relationships among exposure, outcome, and potential reasons for nonparticipation (Question 2 of Table 1). Below, we describe two distinct mechanisms through which such differential participation bias can occur and cause spurious exposure-outcome associations in etiological studies of environmental exposures (Evaluation 2.1 of Table 1). Both mechanisms depend on the fact that only those who survived until the start of follow-up and were healthy enough before enrollment can enroll in the observational cohort studying long-term environmental exposures.

The first example is a frailty-driven mechanism (Figure 1A). The prebaseline health status would determine the probability of participation in the study. Since many environmental exposures are prevalent in the target population long before the start of follow-up and are affecting the probability of surviving or being outcome-free, we would expect to see differential participation bias in such cohorts as long as the environmental exposure has a causal effect on the prebaseline health status (e.g., alive, outcome-free, and healthy enough to participate) and there exists an unmeasured shared risk factor (e.g., unobserved baseline frailty such as respiratory infection) between the prebaseline health status and the probability of having the outcome during follow-up.²⁶ For example, when investigating the effect of air pollution on mortality in a cohort of recruited participants, those who are too sick at baseline will decline to participate in the study. Assuming air pollution,

TABLE 1. Key Points to Consider in Evaluating and Accounting for Differential Participation

Question	Exploratory Analyses/Evaluations
1. What is the target population?	Specified based on the research question
2. Is there differential participation (discrepancy between the cohort and the target population)?	<p>2.1. Draw causal graphs based on subject-matter knowledge and evaluate whether participation creates a backdoor path based on causal graphs.</p> <p>2.2. Evaluate the extent to which the distributions of exposure and outcome in the study cohort deviate from their expected distributions in the target population (can suggest the presence of differential participation but cannot verify the lack of differential participation).</p> <p>2.3. Calculate time-varying effect estimates such as adjusted survival curves and evaluate whether any observed differences in effect estimates over time are consistent with subject-matter knowledge.</p>
3. How to account for bias due to differential participation in effect estimates?	<p>3.1. With additional data from nonparticipants, use inverse probability weighting on participation (or selection) models.</p> <p>3.2. With additional data on all risk factors shared by participation and outcome, control for them using any method.</p> <p>3.3. Without additional data, conduct a washout method that drops the first few years of follow-up. A naïve analysis including all follow-up years might be conducted first to help identify the years to be dropped.</p>

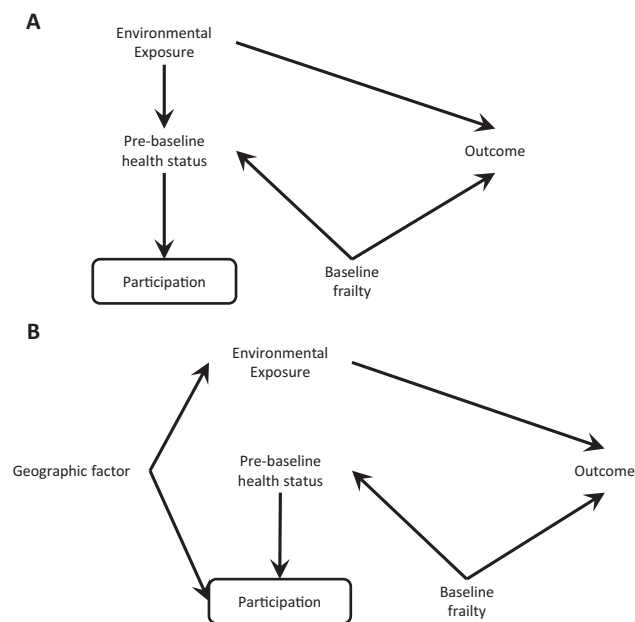


FIGURE 1. Causal graphs for two mechanisms in which differential participation could cause a spurious association between environmental exposure and adverse health outcomes. Bias still exists even if we remove the arrows between exposure and outcome, assuming no direct effect. A, frailty-driven mechanism. B, Geographic factor-driven mechanism.

in combination with other causes of illness, makes people unhealthy to participate in the study, an inverse correlation will emerge between air pollution and mortality through other causes of illness among the participants. That is, among those healthy enough to participate, individuals previously exposed to worse air pollution are less likely to have another unrelated health risk factor such as respiratory infection, and the lack of such risk factor would decrease their chances of mortality.^{27(p100)} The bias arising from selecting participants alive and

outcome-free differs from the bias stemming from selecting participants healthy enough to participate. One could argue that a target population is defined as those free of the outcome at the time of enrollment, and the susceptible population is stable in the target population. Subsequently, one could assert that no differential participation exists in the study population. However, many cohorts also require voluntarily active participation from the subjects, such as completing a questionnaire or attending a medical appointment, which could lead to a stronger connection between exposure and participation through being healthy enough to actively participate at the time of enrollment than simply being outcome-free.²⁸

The second example is a geographic factor-driven mechanism (Figure 1B). Even if the environmental exposure does not affect the prebaseline health status of potential participants, participation could be connected to exposure via their spatial associations with geographical regions. Environmental exposures generally demonstrated spatial heterogeneity across geographical regions. Such geographical regions could be related to participation due to a variety of technical and social reasons. Particularly, cohorts based on existing datasets that were collected for purposes other than studying the etiologic effect of environmental exposure are being increasingly used in studies of environmental exposures, in which a sampling scheme intrinsically involving geographic factors is likely. For example, a survey targeted to study the general population might sample more participants from the urban area for operational reasons, while urbanicity is associated with higher environmental exposures such as air pollution. Such sampling schemes can lead to an open backdoor path between exposure and outcome through participation. This geographic factor-driven bias mechanism would exist even when the environmental exposure of interest is new instead of prevalent.

Although presented separately here, both mechanisms can co-exist in the same cohort. The association between exposure and outcome observed in cohorts affected by either

or both mechanisms could be attenuated or even reversed from the true effect in the target population. Empirical evidence of disparity in exposure and outcome distributions between participants and nonparticipants or between the study population and the target population might suggest differential participation (Evaluation 2.2 of Table 1). However, the lack of disparity in such distributions does not rule out the possibility of differential participation and spurious exposure-outcome association.^{5,29} As a complementary step, estimating time-varying effect estimates like adjusted survival curves could also aid the identification of differential participation as it allows the effect to vary over time, and counterintuitive effects at the beginning of follow-up might suggest differential participation (Evaluation 2.3 of Table 1).

Although differential participation is ideally prevented at the stage of study design, such induced bias could also be controlled via statistical methods in the analytical stage. For example, if we had extra information on those not participating, we could use inverse probability weighting of participation in our analysis to create a pseudo population not affected by differential participation,^{15,30} or use selection models to account for informative missingness (Analysis 3.1 of Table 1).¹¹ If we had information on all risk factors shared by prebaseline health status and outcome, we could also control for them in epidemiologic models to eliminate or mitigate the bias (Analysis 3.2 of Table 1).³¹ However, the above correction methods are infeasible if the required additional data were not collected before the analytical phase, which is very likely. In such settings, where this bias cannot be addressed by considering additional measured covariates, a simple washout method (by removing data from the first few years of follow-up) has been proposed as a possible solution (Analysis 3.3 of Table 1).^{32,33} Some studies, in nonenvironmental settings, have employed a washout analytical method so that the impact of differential participation would decrease as people with high baseline frailty either died or recovered during the washout period and the remaining study population could better approximate the target population.^{34–36} Empirically, the period to be removed could be determined by identifying the elbow or turning point in the adjusted survival curves of the naïve analysis (including all follow-up years), before which the differential participation bias was still strong.³² Below, we demonstrate the effect of differential participation bias in a real-life example, employ the washout analytical method to account for this bias, and conduct a simulation study to demonstrate the efficacy of this method.

Real-life Example: The Canadian Community Health Survey

With a retrospective cohort of respondents to the Canadian Community Health Survey (CCHS), we demonstrate the emergence of counterintuitive associations and the effect of the washout method in accounting for bias. Employing the parametric g-computation demonstrated in Chen et al.³⁷

2023, we aim to evaluate the effectiveness of hypothetical intervention strategies targeting long-term exposure to PM_{2.5} in this real-life example through the comparison of adjusted survival curves with and without the intervention. CCHS is a cross-sectional survey of the general Canadian population for multiple enrolling cycles.³⁸ For illustration, we used data from the 2000/2001 enrolling cycle. We obtained data on the participants' vital statistics, annual exposure to PM_{2.5}, and other time-varying and time-fixed covariates between the survey date (cohort inception) and 31 December 2014 (the end of follow-up) via linkage to administrative datasets and previously estimated exposure surfaces. Several published articles studying environmental exposures used multiple enrollment cycles from this cohort and reported further details.^{39,40} Here we list a few relevant aspects of the cohort (Evaluation 2.1 in Table 1). First, although aimed to study the general population, participation in the CCHS was voluntary and a complete response required participation in an in-person or telephone interview,³⁸ which inherently restricted the participants to be healthier than the general population and might have caused frailty-driven differential participation (Figure 1A). Second, this cohort used health regions to aid sampling and over-sampled rural communities,³⁹ which might cause geographic factor-driven differential participation (Figure 1B).

To explore whether differential participation exists, we plotted age group-specific 4-year cumulative mortality rates over time (Evaluation 2.2 in Table 1). To estimate the adjusted survival curve with and without intervention, we conducted a naïve analysis without considering potential differential participation using 10 years of follow-up data from 2000/2001 to 2010 (Evaluation 2.3 in Table 1). To save computation time, we did not use all 14 years of follow-up in naïve analysis. Specifically, we conducted a parametric g-computation analysis to evaluate the potential health benefits of a hypothetical intervention that reduces participants' long-term exposure to PM_{2.5} to 5 µg/m³ if they were exposed to PM_{2.5} higher than 5 µg/m³, compared to having no intervention. We included a 3-year moving average of PM_{2.5}, indicators for the year, interaction terms between year and PM_{2.5}, rurality, a spline function of age with five knots, and other potential confounders in the pooled logistic model for outcome as part of the parametric g-computation. We included urbanicity as a covariate in the outcome model, and therefore the bias pathway shown in Figure 1B is likely blocked (assuming rurality is the only geographic factor affecting participation). We estimated the 95% confidence intervals of differences in survival probabilities between interventions using standard errors from 200 bootstrap iterations. More details on this analysis were discussed elsewhere.³⁷

Next, as an approach to account for differential participation, we applied the washout method by removing observations in the first 4 years of follow-up by delaying the cohort entry until 2005, with a total of 10 years of follow-up data until

2014, followed by repeating the parametric g-computation analysis (Analysis 3.3 in Table 1). We incorporated four more years of follow-up than those utilized in the naïve analysis so that the length of follow-up would be comparable. We decided to delay the cohort entry until 2005 because the difference in survival probability reached the inflection point in the fifth year in the naïve analysis (Figure 2A), suggesting that the bias caused by differential selection was weakened from that time point onwards. In addition, we excluded those older than 79 in 2005 in the washout analysis to ensure that all participants would be followed for up to 10 years or until death because the original administrative dataset censored individuals after age 89.

Last, to facilitate comparison with traditional survival analysis methods, we repeated the naïve and washout analyses using a Cox proportional hazards model and an Aalen additive hazard model with 3, 5, and 10 years of follow-up. Like previous studies using traditional survival analysis methods, we assumed a constant association between $PM_{2.5}$ and mortality and included the same set of confounders as those used in the outcome model of parametric g-computation other than indicators for the year and interaction terms between year and $PM_{2.5}$. All analyses of the CCHS cohort were carried out in R version 4.0.5⁴¹ and relevant codes for g-computation can be found at: https://github.com/suthlam/cchs_g_computation.git. The Health Canada-Public Health Agency of Canada Research Ethics Board approved the study.

Simulation Scenarios

For illustrative purposes and to demonstrate that the two hypothesized bias mechanisms discussed above could plausibly attenuate or even reverse the true effect, we conducted simulations for each mechanism separately based on causal graphs in eFigure 1; <http://links.lww.com/EDE/C103>.

The underlying structures of bias are the same as shown in Figure 1 except that we added the direct effect from exposure to the outcome and updated the label of the knot to mimic the CCHS example for easier interpretation.

We generated time-to-event data using structural equations of additive hazards with modified simulation codes from Strohmaier et al. 2015.⁴² To mimic the real world, we used coefficients based on statistics and effect estimates from the CCHS cohort (e.g., 0.0002 increase in hazard rate per $1 \mu\text{g}/\text{m}^3$ increase in $PM_{2.5}$ for the association between $PM_{2.5}$ and mortality). For simplicity and didactic purposes, we used time-fixed exposure and assumed no direct effect from baseline frailty (frailty-driven mechanism) or prebaseline health status (geographic factor-driven mechanism) on death after 3 years of follow-up. We also assumed no unmeasured confounding between residential $PM_{2.5}$ and mortality. The simulation involved three steps: (1) we simulated a full cohort of 100,000 individuals with 10 years of follow-up and no differential participation; (2) we created an observed cohort by only including 70% of individuals with the highest probability of participating at baseline, which is affected by a geographic factor (geographic factor-driven mechanism only) and the prebaseline health status (both mechanisms) separately; and (3) we iterated each simulation 100 times. Specific coefficients and distributions used in the simulation are included in eAppendix 1; <http://links.lww.com/EDE/C103>.

To estimate the association between $PM_{2.5}$ and death, we repeated the same analyses for the simulated full cohort (representing the target population) and observed cohort (representing the analytical cohort) separately as we did for the CCHS cohort: g-computation, Aalen model, and Cox model with and without applying the washout method. The maximum length of follow-up year is 10 in naïve analysis and seven in washout analysis. Details are included in

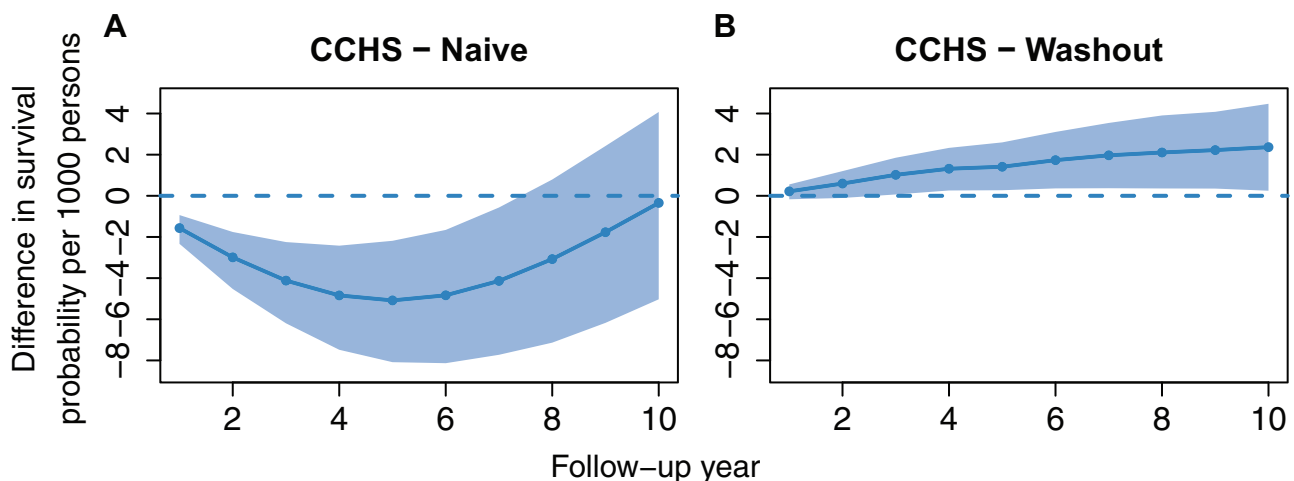


FIGURE 2. Difference in survival probability ($5 \mu\text{g}/\text{m}^3$ threshold intervention minus no intervention) over time estimated with parametric g-computation in the CCHS cohort. A, naïve analysis in using 10 years of follow-up. B, Washout analysis where the first 4 or 5 years of follow-up are dropped. Shaded bands are 95% confidence intervals estimated with bootstrapping.

eAppendix 2; <http://links.lww.com/EDE/C103>. Aside from visually comparing summaries of effect estimates from the full and observed cohorts, we also calculated absolute bias as the difference between effect estimates in the observed cohort and the corresponding full cohort before summarizing the bias across iterations. This comparison used estimates from the full cohort as the target parameters and reduced the influence of random errors generated by the data simulation process. We report absolute biases from the Aalen model and relative biases from the Cox model (percentage difference in hazard ratios between the observed and the full cohort). 95% simulation intervals (SIs) were calculated using the Wald standard error of estimates from the 100 iterations. All simulations and analyses were carried out in R 4.1.0 and relevant codes can be found at: https://github.com/suthlam/differential_participation_simulation.git.

RESULTS

The naïve analysis of the CCHS cohort included a final cohort of 65,470 individuals. Using parametric g-computation, we found negative values for the difference in annual probabilities of survival (with the 5 µg/m³ threshold intervention minus without intervention) during most years of follow-up

(Figure 2A), suggesting a worse survival probability in the study population after reducing the level of ambient PM_{2.5}. This finding is contrary to subject-matter knowledge about the harmful effect of chronic exposure to PM_{2.5} on mortality.⁴³ Additionally, we observed low cumulative mortality rates in the first few years of follow-up across all age groups (eFigure 2; <http://links.lww.com/EDE/C103>). Results from the Aalen model and the Cox model were consistent with those from the g-computation, with null or negative associations between residential PM_{2.5} and mortality across all follow-up periods (Table 2, eFigure 3A and C; <http://links.lww.com/EDE/C103>).

After applying the washout method to the CCHS cohort, the cohort size was reduced to 62,365. The estimated differences in annual probabilities of survival were positive during 10 years of follow-up after applying washout analysis, suggesting a beneficial effect of PM_{2.5} reduction on mortality (Figure 2B). We also observed positive associations for all follow-up periods in the Aalen model (Table 2, eFigure 3A; <http://links.lww.com/EDE/C103>), while the associations were null in the Cox model (Table 2, eFigure 3C; <http://links.lww.com/EDE/C103>).

Using simulations with the frailty-driven and geographic factor-driven mechanisms, we successfully mimicked the patterns of estimates observed in the naïve analysis of the CCHS cohort in the observed cohort simulated. Like

TABLE 2. Effect Estimates When Differential Participation Exists With and Without Applying the Washout Method in Analysis by Cohort, Model, and Length of Follow-up Time

Cohort	Analysis	Length of Follow-up Time	Hazard Difference Per Unit Change in PM _{2.5} per 1000 Persons (95% CI for CCHS, 95% SI for Simulated Cohorts) (Aalen model)		Hazard Ratio Per Unit Change in PM _{2.5} (95% CI for CCHS, 95% SI for Simulated Cohorts) (Cox Model)	
			Full Cohort	Observed Cohort	Full Cohort	Observed Cohort
CCHS	Naïve	3-year	NA	0.03 (−0.15, 0.22)	NA	0.99 (0.96, 1.03)
		5-year		−0.18 (−0.36, −0.01)		0.98 (0.96, 1.00)
		10-year		−0.13 (−0.29, 0.02)		0.99 (0.98, 1.01)
	Washout	3-year		0.40 (0.10, 0.70)		0.99 (0.96, 1.03)
		5-year		0.63 (0.38, 0.87)		1.01 (0.98, 1.04)
		10-year ^a		0.38 (0.19, 0.58)		1.01 (0.99, 1.03)
Frailty-driven simulation	Naïve	3-year	0.12 (−0.06, 0.29)	−0.05 (−0.26, 0.17)	1.02 (1.00, 1.04)	0.99 (0.97, 1.02)
		5-year	0.15 (0.03, 0.27)	0.01 (−0.16, 0.17)	1.02 (1.01, 1.04)	1.00 (0.98, 1.03)
		10-year	0.18 (0.09, 0.27)	0.11 (−0.01, 0.22)	1.02 (1.01, 1.04)	1.02 (1.00, 1.03)
	Washout	3-year	0.13 (−0.04, 0.30)	0.14 (−0.07, 0.35)	1.03 (1.00, 1.05)	1.03 (0.99, 1.06)
		5-year	0.16 (0.02, 0.29)	0.16 (−0.02, 0.35)	1.03 (1.01, 1.04)	1.03 (1.00, 1.05)
		7-year ^b	0.17 (0.07, 0.28)	0.18 (0.03, 0.32)	1.03 (1.01, 1.04)	1.03 (1.01, 1.05)
Geographic factor-driven simulation	Naïve	3-year	0.12 (−0.01, 0.25)	−0.01 (−0.16, 0.13)	1.02 (1.01, 1.04)	1.00 (0.98, 1.02)
		5-year	0.16 (0.07, 0.25)	0.03 (−0.08, 0.15)	1.02 (1.01, 1.04)	1.01 (1.00, 1.02)
		10-year	0.18 (0.11, 0.25)	0.12 (0.04, 0.2)	1.03 (1.02, 1.03)	1.02 (1.01, 1.03)
	Washout	3-year	0.14 (0.02, 0.26)	0.13 (−0.01, 0.27)	1.03 (1.01, 1.05)	1.03 (1.01, 1.05)
		5-year	0.16 (0.06, 0.27)	0.17 (0.05, 0.28)	1.03 (1.01, 1.04)	1.03 (1.01, 1.05)
		7-year ^b	0.17 (0.09, 0.26)	0.18 (0.07, 0.29)	1.03 (1.02, 1.04)	1.03 (1.01, 1.04)

^aIn the real-life example of CCHS, we have 14 years of follow-up time thus it is possible to have 10 years of follow-up time after dropping the first 4 or 5 years of follow-up in the washout analyses.

^bIn the simulation, we created a cohort with 10 years of follow-up time thus we only have 7 years of follow-up time after dropping the first 3 years of follow-up in the washout analyses.

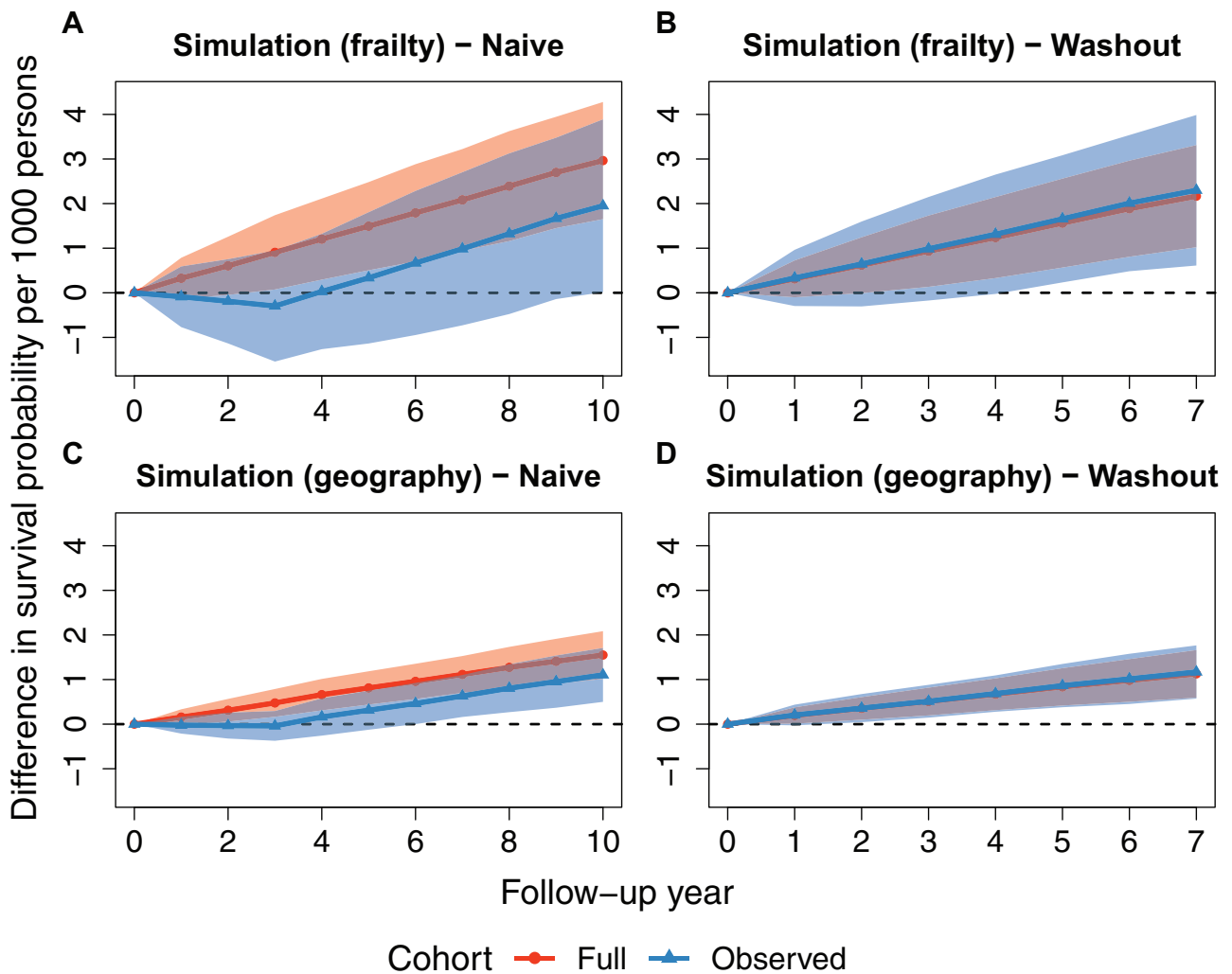


FIGURE 3. Difference in survival probability ($5 \mu\text{g}/\text{m}^3$ threshold intervention minus no intervention) over time estimated with parametric g-computation in simulated cohorts. A, naïve analysis in cohorts under frailty-driven mechanism using 10 years of follow-up; B, washout analysis in cohorts under frailty-driven mechanism where first 3 years of follow-up are dropped; C & D, naïve and washout analyses in cohorts under geographic factor-driven mechanism. Shadowed bands are 95% confidence intervals estimated with bootstrapping. In simulation results, the cumulative survival probabilities were calculated by standardizing to confounder distributions of the observed cohort (see eAppendix 2; <http://links.lww.com/EDE/C103> for more details).

g-computation results from the CCHS cohort, we observed negative differences in survival probability comparing $5 \mu\text{g}/\text{m}^3$ threshold intervention to no intervention under both bias mechanisms in the observed cohorts from the simulation without applying the washout method, and the differences became positive in later years of follow-up (Figure 3A,C). The differences in survival probability in the full cohort were positive (Figure 3A,C). In analyses of simulations using Aalen and Cox models, we found null or negative associations between $\text{PM}_{2.5}$ and mortality in most simulations of the observed cohort with shorter follow-up periods in analysis, while the associations were positive in most simulations of the full cohort (Table 2, eFigure 3B and 3D; <http://links.lww.com/EDE/C103>). When directly comparing estimates of the

observed and full cohort from one iteration of the simulation, we found negative bias for both Aalen and Cox models (Figure 4, eTable 1; <http://links.lww.com/EDE/C103>). The bias became smaller as we included more follow-up years in the analysis, suggesting that a long follow-up period dilutes the initial period of severe bias.

With the washout method applied to the simulated cohorts, we observed positive differences in survival probability under both bias mechanisms for the observed cohorts, which were close to the corresponding estimates for the full cohorts (Figure 3B,3D). Furthermore, the bias became negligible when washout analysis was applied in the Aalen model and the Cox model (Figure 4, eTable 1; <http://links.lww.com/EDE/C103>).

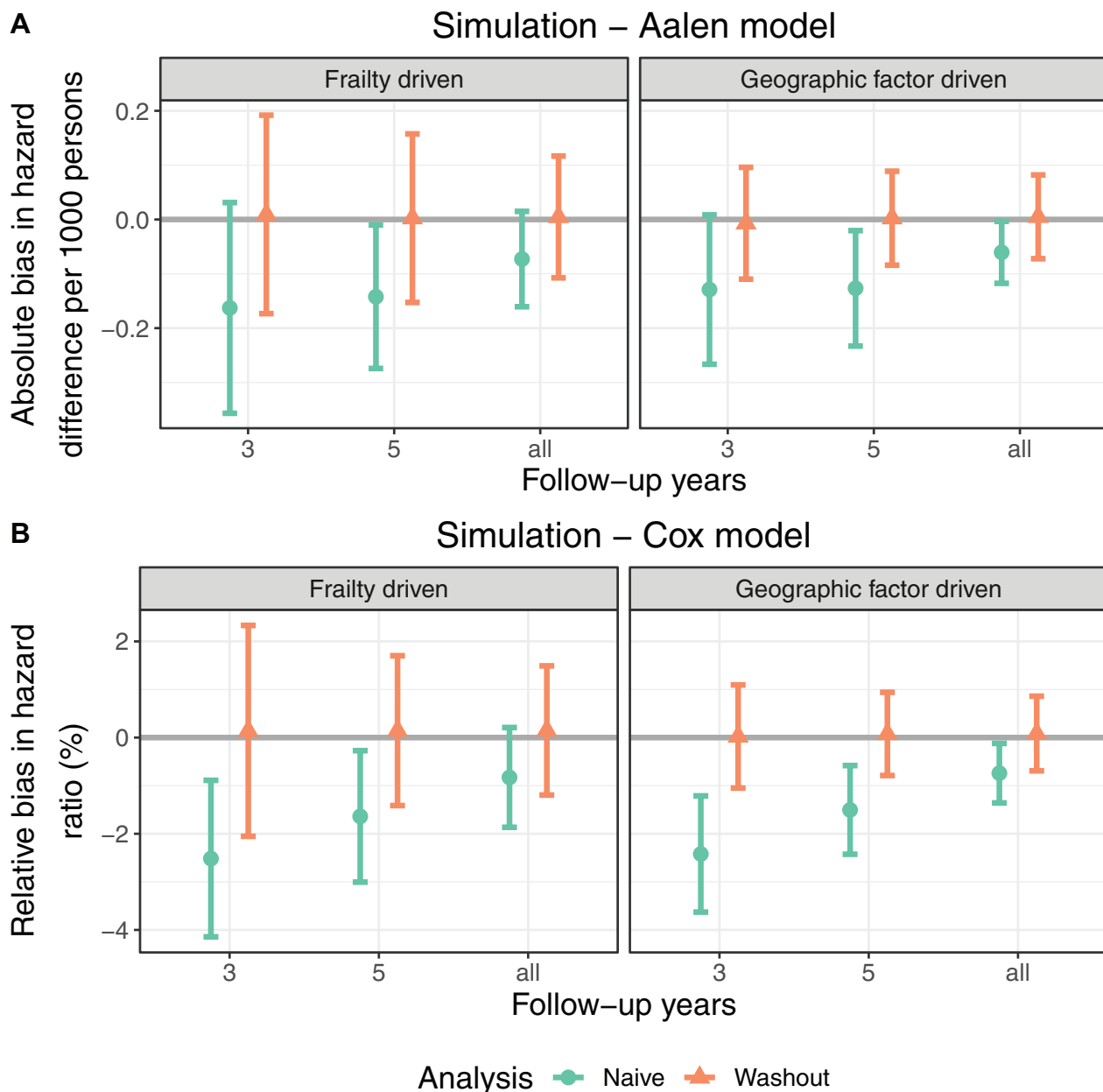


FIGURE 4. Absolute bias from the Aalen model (effect estimate per unit change in $PM_{2.5}$ in the simulated observed cohort minus the effect estimate in the simulated full cohort) and relative bias from the Cox model (absolute bias divided by the effect estimate in the simulated full cohort) when differential participation exists, with and without applying washout in analysis by bias mechanism and follow-up time. “All” follow-up year represents a 10-year follow-up for the naive analysis and a 7-year follow-up for the washout analysis.

DISCUSSION

As large cohorts not originally designed to study environmental exposures are being increasingly used to answer etiological questions in environmental research, spurious or even counterintuitive associations are more likely to occur due to differential participation. It is important to be mindful of the possibility of bias from differential participation in such

settings. In this article, we discussed key points to consider in evaluating and accounting for differential participation (Table 1); we described two distinct mechanisms of differential participation that might cause counterintuitive associations in observational cohorts of environmental exposures using causal graphs (Figure 1). We demonstrated how differential participation due to the selection of healthier individuals into

the cohort could lead to a counterintuitive protective association between long-term air pollution and mortality in a real-life example based on the CCHS cohort. We also successfully mimicked patterns in the CCHS cohort using simulation, which confirmed that the proposed mechanisms could plausibly cause counterintuitive associations. Last, we described the washout analysis, removing data from the first few years of follow-up, as one viable analytical solution to differential participation that does not require additional data.

Although both proposed bias mechanisms for differential participation might happen simultaneously in the same cohort studying long-term environmental exposures, different cohorts might be more susceptible to one or the other. Both bias structures rely on the fact that the prebaseline health status (e.g., death/outcome of interest or illness) of people in the target population would likely affect their probability of participation in the cohorts. The bias through death or experiencing the outcome of interest before cohort initiation (i.e., left-truncation) could also be framed as a failure of target trial emulation due to misalignment in the start of exposure and the start of follow-up, which is inherent to studies of long-term/prevalent exposure.⁴⁴ If we have a birth cohort with many decades of follow-up, using a target trial emulation framework, we could align the start of eligibility, exposure, and follow-up in our study cohort and estimate the effect of exposure duration or trajectory from the time of birth.⁴⁵ Alternatively, if we are studying a cohort where a stable susceptible population could be assumed, the impact of this part of the bias might be negligible. If not, we might consider exploring a different type of question by treating previous exposure as a confounder and evaluating the impact of change in exposure (i.e., treatment decision design).^{46,47}

On the other hand, the bias through illness (affecting participation but not necessarily related to outcome) might be more severe because it is more common than mortality. Thus, cohorts that require extra voluntary in-person activities such as filling out questionnaires or attending medical appointments could experience a more severe nonrepresentativeness. Examples of such cohorts include the UK Biobank cohort,^{48–50} the CCHS,^{39,51} and the US National Institutes of Health–AARP Diet & Health cohort.⁵² Similarly, cohorts with a higher probability of losing participants before enrollment might experience a more severe nonrepresentativeness, such as cohorts restricted to live births¹⁶ and existing employees.⁴ Besides, the geographic factor-driven mechanism assumes that some geographic factors are directly associated with the exposure of interest and affect participation. Cohorts not originally designed to explore the etiological impacts of environmental exposures are particularly prone to this bias structure because the sampling scheme in the original design is likely driven by geographic factors such as rurality for operational reasons. These mechanisms also apply to cohorts based on electronic health records, in which case it is healthcare

utilization, instead of prebaseline health status, that affects participation in the cohort.

Calculation of time-specific effect estimates, such as adjusted survival curves, is essential to identify the differential participation bias, while a single estimate of association averaged across the study period (e.g., a hazard ratio from the Cox model or a hazard difference from the Aalen model) could hide the heterogeneity in effect estimates over time. As the spurious association between exposure and outcome caused by differential participation would gradually disappear over time (when the follow-up period is long enough), a single estimate of association averaged across a long enough period would only seem to be slightly attenuated compared with what is expected in the target population. Such issues regarding the use of hazard ratios have been previously described,^{53,54} and many authors recommended relying on alternative methods based on discrete-time models so that time-specific effect estimates can be obtained. When estimating time-specific effect estimates with adjusted survival curves (Figure 2A), we were able to detect such bias in the earliest periods and attempt to account for such bias using the washout method (Figure 2B). In addition, adjusted survival curves are more causally interpretable compared to single-effect estimates such as the hazard ratio.⁵³ Adjusted survival curves could also help choose the appropriate washout period.

Although results in both the real-life example and simulations supported the efficacy of the washout method in achieving unbiased estimates when differential participation exists, it is an *ad hoc* method with limitations. The washout method aims to start analysis when the arrows connecting prebaseline health status and the outcome (i.e., death) in later times become weaker as those frailer participants with lower exposure died over time or the participant's health status changed over time, which could better approximate the target population. Although our simulation demonstrated an elimination of bias after applying the washout method, its success depends on two conditions: (1) correctly excluding enough of the follow-up time; and (2) the effect of the exposure is relatively constant over time (i.e., the susceptible population remained steady). In a real-life scenario, we could estimate the length of time to exclude by identifying the turning point in the adjusted survival curve of the naïve analysis. However, the connection between differential participation and outcome during later times of follow-up might decrease gradually with no clear turning point, making the decision a bias-variance tradeoff with no clear inflection point to choose. In that case, it is unlikely that the washout method could remove the bias entirely. Alternative analytical methods such as inverse probability weighting (for participation or censoring) requiring additional data could eliminate the bias in both mechanisms by removing arrows in the bias pathway and should be considered if a sufficient set of participation predictors were measured.^{15,55}

Aside from collider stratification bias, selection bias could also be caused by conditioning on an effect modifier of the exposure-outcome association without being a collider (i.e., differential participation due to an effect modifier). Such mechanisms are unlikely to cause a counterintuitive association because the observed association is biased for the target population but unbiased for the study population.⁵⁶ Nonetheless, they can still lead to invalid inferences for the target population.²⁹ It is also worth mentioning that other paths of selection bias might lead to spurious associations in observational cohorts, including competing risk/right censoring,^{10,57–59} which warrants future discussion in the context of observational cohorts studying environmental exposures.

Finally, although differential participation could lead to spurious associations in observational cohorts studying environmental exposures, we are not suggesting that researchers should systematically apply the washout analysis to all observational cohorts. Instead, we would like to draw attention to this differential participation pattern (and different mechanisms driving it) so that environmental epidemiologists can investigate such patterns in their observational cohorts, especially when counterintuitive associations are observed. Furthermore, the bias introduced by differential participation might be disguised in single estimates of association averaged across a relatively long study period (e.g., hazard ratio). Therefore, we recommend systematically conducting analyses (at least as sensitivity analyses) that derive time-specific effect estimates to identify not only such differential participation but also any potential violations of the proportional hazards assumption when using Cox models.

REFERENCES

- Lash T, VanderWeele T, Haneuse S, Rothman K. *Modern Epidemiology*. 4th ed. Wolters Kluwer Health; 2020.
- Rothman KJ, Gallacher JEJ, Hatch EE. Why representativeness should be avoided. *Int J Epidemiol*. 2013;42:1012–1014.
- Munafò MR, Tilling K, Taylor AE, Evans DM, Davey Smith G. Collider scope: when selection bias can substantially influence observed associations. *Int J Epidemiol*. 2018;47:226–235.
- Applebaum KM, Malloy EJ, Eisen EA. Left truncation, susceptibility, and bias in occupational cohort studies. *Epidemiology*. 2011;22:599–606.
- Greenland S. Response and follow-up bias in cohort studies. *Am J Epidemiol*. 1977;106:184–187.
- Howards PP, Hertz-Picciotto I, Poole C. Conditions for bias from differential left truncation. *Am J Epidemiol*. 2007;165:444–452.
- Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:11.
- Cain KC, Harlow SD, Little RJ, et al. Bias due to left truncation and left censoring in longitudinal studies of developmental and disease processes. *Am J Epidemiol*. 2011;173:1078–1084.
- Hazelbag CM, Klungel OH, van Staa TP, de Boer A, Groenwold RHH. Left truncation results in substantial bias of the relation between time-dependent exposures and adverse events. *Ann Epidemiol*. 2015;25:590–596.
- López-López JA, Sterne JAC, Higgins JPT. Selection bias introduced by informative censoring in studies examining effects of vaccination in infancy. *Int J Epidemiol*. 2019;48:2001–2009.
- Muñoz J, Hufstедler H, Gustafson P, Bärnighausen T, De Jong VMT, Debray TPA. Dealing with missing data using the Heckman selection model: methods primer for epidemiologists. *Int J Epidemiol*. 2023;52:5–13.
- Huang JY. Representativeness is not representative: addressing major inferential threats in the UK biobank and other big data repositories. *Epidemiology*. 2021;32:189–193.
- Cheung KL, ten Klooster PM, Smit C, de Vries H, Pieterse ME. The impact of non-response bias due to sampling in public health studies: a comparison of voluntary versus mandatory recruitment in a Dutch national survey on adolescent health. *BMC Public Health*. 2017;17:276.
- Richiardi L, Boffetta P, Merletti F. Analysis of nonresponse bias in a population-based case-control study on lung cancer. *J Clin Epidemiol*. 2002;55:1033–1040.
- Danaei G, García Rodríguez LA, Fernandez Cantero O, Hernán MA. Statins and risk of diabetes: an analysis of electronic medical records to evaluate possible bias due to differential survival. *Diabetes Care*. 2013;36:1236–1240.
- Raz R, Kioumourtzoglou MA, Weisskopf MG. Live-birth bias and observed associations between air pollution and autism. *Am J Epidemiol*. 2018;187:2292–2296.
- Dahabreh IJ, Kent DM. Index event bias: an explanation for the paradoxes of recurrence risk research. *JAMA*. 2011;305:822–823.
- Stensrud MJ, Valberg M, Røysland K, Aalen OO. Exploring selection bias by causal frailty models: the magnitude matters. *Epidemiology*. 2017;28:379–386.
- Weiskopf NG, Dorr DA, Jackson C, Lehmann HP, Thompson CA. Healthcare utilization is a collider: an introduction to collider bias in EHR data reuse. *J Am Med Inform Assoc*. 2023;30:ocad013.
- Banack HR, Kaufman JS. The obesity paradox: understanding the effect of obesity on mortality among individuals with cardiovascular disease. *Prev Med*. 2014;62:96–102.
- Lajous M, Banack HR, Kaufman JS, Hernán MA. Should patients with chronic disease be told to gain weight? the obesity paradox and selection bias. *Am J Med*. 2015;128:334–336.
- Lisonkova S, Joseph KS. Left truncation bias as a potential explanation for the protective effect of smoking on preeclampsia. *Epidemiology*. 2015;26:436–440.
- Rodríguez-López M, Escobar MF, Merlo J, Kaufman JS. Reevaluating the protective effect of smoking on preeclampsia risk through the lens of bias. *J Hum Hypertens*. 2023;37:338–344.
- Leung M, Kioumourtzoglou MA, Raz R, Weisskopf MG. Bias due to selection on live births in studies of environmental exposures during pregnancy: a simulation study. *Environ Health Perspect*. 2021;129:47001.
- Kinlaw AC, Buckley JP, Engel SM, Poole C, Brookhart MA, Keil AP. Left truncation bias to explain the protective effect of smoking on preeclampsia: potential, but how plausible? *Epidemiology*. 2017;28:428–434.
- Flanders WD, Eldridge RC, McClellan W. A nearly unavoidable mechanism for collider bias with index-event studies. *Epidemiology*. 2014;25:762–764.
- Hernán MA, Robins JM. *Causal Inference: What If*. Published online 2020:311.
- Cohen G, Duffy JC. *Are Nonrespondents to Health Surveys Less Healthy Than Respondents?* Published online 2002:11.
- Lu H, Cole SR, Howe CJ, Westreich D. Toward a clearer definition of selection bias when estimating causal effects. *Epidemiology*. 2022;33:699–706.
- Reeves A, Elliott MR, Lewis TT, Karvonen-Gutierrez CA, Herman WH, Harlow SD. Study selection bias and racial or ethnic disparities in estimated age at onset of cardiometabolic disease among midlife women in the US. *JAMA Netw Open*. 2022;5:e2240665.
- VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics*. 2011;67:1406–1413.
- Törner A, Duberg AS, Dickman P, Svensson A. A proposed method to adjust for selection bias in cohort studies. *Am J Epidemiol*. 2010;171:602–608.
- Törner A, Dickman P, Duberg AS, et al. A method to visualize and adjust for selection bias in prevalent cohort studies. *Am J Epidemiol*. 2011;174:969–976.
- Lee IM, Manson JE, Hennekens CH, Paffenbarger RS. Body weight and mortality a 27-year follow-up of middle-aged men. *JAMA*. 1993;270:2823–2828.

35. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol*. 2009;38:1599–1611.
36. Bohrmann B, Massa MS, Ross S, Lewington S, Lacey B. Body mass index and risk of hospitalization or death due to lower or upper respiratory tract infection. *JAMA*. 2023;329:1512–1514.
37. Chen C, Chen H, van Donkelaar A, et al. Using parametric g-computation to estimate the effect of long-term exposure to air pollution on mortality risk and simulate the benefits of hypothetical policies: the Canadian Community Health Survey cohort (2005 to 2015). *Environ Health Perspect*. 2023;131:37010.
38. Government of Canada SC. *Canadian Community Health Survey (CCHS) Cycle 1.1*. Published October 24, 2007. Available at: <https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&Id=3359>. Accessed 26 September 2021.
39. Christidis T, Erickson AC, Pappin AJ, et al. Low concentrations of fine particle air pollution and mortality in the Canadian community health survey cohort. *Environ Health*. 2019;18:84.
40. Pinault L, Tjepkema M, Crouse DL, et al. Risk estimates of mortality attributed to low concentrations of ambient fine particulate matter in the Canadian community health survey cohort. *Environ Health*. 2016;15:18.
41. R Core Team. *R: A language and environment for statistical computing*. Published online 2021. Available at: <https://www.R-project.org/>
42. Strohmaier S, Røysland K, Hoff R, Borgan O, Pedersen TR, Aalen OO. Dynamic path analysis – a useful tool to investigate mediation processes in clinical survival trials. *Stat Med*. 2015;34:3866–3887.
43. Chen J, Hoek G. Long-term exposure to PM and all-cause and cause-specific mortality: a systematic review and meta-analysis. *Environ Int*. 2020;143:105974.
44. Hernán MA, Sauer BC, Hernández-Díaz S, Platt R, Shrier I. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–75.
45. Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ*. 2018;360:k182.
46. Brookhart MA. Counterpoint: the treatment decision design. *Am J Epidemiol*. 2015;182:840–845.
47. Neophytou AM, Costello S, Picciotto S, et al. Accelerated lung function decline in an aluminum manufacturing industry cohort exposed to PM_{2.5}: an application of the parametric g-formula. *Occup Environ Med*. 2019;76:888–894.
48. Aung N, Sanghvi MM, Zemrak F, et al. Association between ambient air pollution and cardiac morpho-functional phenotypes: insights from the UK Biobank population Imaging Study. *Circulation*. 2018;138:2175–2186.
49. Wang M, Zhou T, Song Q, et al. Ambient air pollution, healthy diet and vegetable intakes, and mortality: a prospective UK Biobank study. *Int J Epidemiol*. 2022;51:1243–1253.
50. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault JB, Kutalik Z. Participation bias in the UK biobank distorts genetic associations and downstream analyses. *Nat Hum Behav*. 2023;7:1216–1227.
51. Pappin AJ, Christidis T, Pinault LL, et al. Examining the shape of the association between low levels of fine particulate matter and mortality across three cycles of the Canadian census health and environment cohort. *Environ Health Perspect*. 2019;127:107008.
52. Thurston GD, Ahn J, Cromar KR, et al. Ambient particulate matter air pollution exposure and mortality in the NIH-AARP diet and health cohort. *Environ Health Perspect*. 2016;124:484–490.
53. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21:13–15.
54. Stensrud MJ, Hernán MA. Why test for proportional hazards? *JAMA*. 2020;323:1401–1402.
55. Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. *Acta Obstet Gynecol Scand*. 2018;97:407–416.
56. Hernán MA. Invited commentary: selection bias without colliders. *Am J Epidemiol*. 2017;185:1048–1050.
57. Ozenne BMH, Scheike TH, Stærk L, Gerds TA. On the estimation of average treatment effects with right-censored time to event outcome and competing risks. *Biom J*. 2020;62:751–763.
58. Rudolph JE, Lesko CR, Naimi AI. Causal inference in the face of competing events. *Curr Epidemiol Rep*. 2020;7:125–131.
59. Schuster NA, Hoogendijk EO, Kok AAL, Twisk JWR, Heymans MW. Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. *J Clin Epidemiol*. 2020;122:42–48.