American Journal of Epidemiology

Vol. 192, No. 10
https://doi.org/10.1093/aje/kwad135
Advance Access publication:
July 3, 2023

Practice of Epidemiology

Immortal Time Bias With Time-Varying Exposures in Environmental Epidemiology: A Case Study in Lung Cancer Survival

Paige Sheridan*, Chen Chen, Caroline A. Thompson, and Tarik Benmarhnia

* Correspondence to Dr. Paige Sheridan, Herbert Wertheim School of Public Health, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093 (e-mail: paigesheridan23@gmail.com).

Initially submitted February 22, 2022; accepted for publication June 4, 2023.

Immortal time bias is a well-recognized bias in clinical epidemiology but is rarely discussed in environmental epidemiology. Under the target trial framework, this bias is formally conceptualized as a misalignment between the start of study follow-up (time 0) and treatment assignment. This misalignment can occur when attained duration of follow-up is encoded into treatment assignment using minimums, maximums, or averages. The bias can be exacerbated in the presence of time trends commonly found in environmental exposures. Using lung cancer cases from the California Cancer Registry (2000–2010) linked with estimated concentrations of particulate matter less than or equal to 2.5 μm in aerodynamic diameter ($PM_{2.5}$), we replicated previous studies that averaged $PM_{2.5}$ exposure over follow-up in a time-to-event model. We compared this approach with one that ensures alignment between time 0 and treatment assignment, a discrete-time approach. In the former approach, the estimated overall hazard ratio for a 5-$\mu g/m^3$ increase in $PM_{2.5}$ was 1.38 (95% confidence interval: 1.36, 1.40). Under the discrete-time approach, the estimated pooled odds ratio was 0.99 (95% confidence interval: 0.98, 1.00). We conclude that the strong estimated effect in the former approach was likely driven by immortal time bias, due to misalignment at time 0. Our findings highlight the importance of appropriately conceptualizing a time-varying environmental exposure under the target trial framework to avoid introducing preventable systematic errors.

cancer; causal inference; environmental epidemiology; immortal time bias; observational studies; target trials

Abbreviations: CCR, California Cancer Registry; CI, confidence interval; HR, hazard ratio; ICDO-3, *International Classification of Diseases for Oncology, Third Edition*; ITB, immortal time bias; $PM_{2.5}$, particulate matter less than or equal to 2.5 μm in aerodynamic diameter.

Epidemiologic studies often seek to measure the impact of time-varying environmental exposures on time-to-event outcomes. In this setting, it is necessary to appropriately consider the time-varying nature of the exposure. However, researchers often overlook this requirement and attempt to simplify the analysis by treating the exposure as time-fixed to avoid the added complexity of accounting for a time-varying exposure. For instance, some investigators choose to average exposure over follow-up for each individual (1–5). While this may initially appear to be an appropriate simplification, assigning baseline exposure using an averaged exposure can introduce immortal time bias (ITB) into the study. Described generally, ITB occurs when there is a misalignment between time 0 and treatment assignment. In this paper, we will use the target trial framework to identify how averaging a time-varying environmental exposure in a time-to-event analysis can introduce ITB into an observational study. We then illustrate an approachable analytical solution that prevents ITB and appropriately considers the time-varying nature of these exposures. The terms "exposure" and "treatment" will be used synonymously throughout this paper.

The target trial framework views an observational study through the lens of a hypothetical randomized trial (target trial) to avoid fundamental flaws in the analysis that can result in biases, including ITB (6). This framework specifies that the fundamental principles that guide the design and analysis of randomized trials must be applied in observational studies to avoid systematic errors. One of these central principles is the alignment of start of follow-up (time 0),

eligibility criteria, and treatment assignment. This alignment at time 0 occurs by design in a randomized trial (at the time of randomization), so when this principle is ignored in an observational study, systematic errors can be introduced. One of the errors that can occur if this principle of alignment is violated is "classical" ITB. Here, ITB occurs when the start of follow-up and eligibility criteria are aligned but treatment assignment is made using information that occurs after baseline (i.e., "future" looking) (7), which creates a misalignment between treatment assignment and time 0 (see Web Figure 1, available at https://doi.org/10.1093/aje/kwad135). For the purposes of this discussion, we will refer to the "principle of alignment at time 0" to describe the alignment between time 0 and treatment assignment (assuming alignment of eligibility criteria).

While ITB is most often identified when there is a well-defined period of "immortal time," as first described by Gail (8) in the context of heart transplantation studies in 1972 (see details in the Web Appendix), there are other subtle ways postbaseline information may be used to assign treatment status. One example of this occurs when treatment assignment is dependent on an individual's attained duration of follow-up (9). This appears in studies that use maximums, minimums, or averages of exposure during follow-up as the exposure of interest (7). This has been described in studies of "long-term" and "short-term" drug use, where a participant must have had a longer duration of follow-up to be classified as a long-term user (10, 11). For example, in an observational study that demonstrated a protective effect of statin use on lung cancer risk when comparing long-term statin users with nonusers (12), individuals were classified using their attained duration of statin therapy *during* follow-up. While there is no obvious period of "immortal time" here, we can identify that the principle of alignment at time 0 is violated as a result of assigning treatment using attained duration of follow-up, resulting in ITB.

Because ITB has most commonly been addressed in clinical epidemiology and pharmacoepidemiology (13–15), its occurrence in environmental epidemiology has been largely overlooked. However, using the target trial framework, we can identify that ITB can be introduced when time-varying environmental exposures are averaged over follow-up. Averaging an exposure over the course of follow-up defines the exposure value based on an individual's attained length of follow-up, creating misalignment of treatment and time 0 (as described above). This results in a study sample in which individuals with longer follow-up times will have more exposure observations included in their average than individuals with shorter follow-up times (see Web Figure 2). These averaged exposures (between individuals with different follow-up times) are no longer inherently comparable as a result of the different number of values included in the calculation. Conceptually, another way in which this error in exposure assignment can be understood is that using an average exposure over follow-up as a baseline exposure inappropriately attributes exposure to a time period where that exposure has not yet accrued.

While we can identify that using a postbaseline averaged exposure can introduce ITB under the target trial framework, the induced systematic error can be even greater when the exposure has a strong temporal trend, as with air pollution (16, 17). When this is the case, the averaged exposure level is associated not only with length of follow-up but also with the temporal trend. For example, if an exposure is decreasing over calendar time, participants with longer follow-up times will have lower averages than those with shorter follow-up times, due to the temporal trend in the exposure. This results in an even greater discrepancy in assigned exposure values between those with longer and shorter follow-up times, which amplifies the magnitude of such bias.

To date, no studies (to our knowledge) have described how averaging time-varying environmental exposures over follow-up time in a time-to-event analysis can introduce ITB. In this paper, we present a case study examining the impact of air pollution on 5-year survival after a lung cancer diagnosis. We use 2 approaches to estimate this association. First, we define air pollution exposure using the average exposure over follow-up (we refer to this as the naive approach). We then use an alternate method, a discrete-time approach, that appropriately assigns air pollution exposure at the time it has been accrued, and in turn upholds the principle of alignment at time 0. We then compare effect estimates between the naive and discrete-time approaches to illustrate how ITB can impact effect estimates.

CASE STUDY

Several studies have examined the impact of air pollution exposure after a cancer diagnosis on cancer survival in a time-to-event analysis (1–5). These authors treated air pollution exposure as time-fixed by averaging air pollution exposure from the date of diagnosis to the date of last follow-up and used Cox proportional hazards models to estimate hazard ratios (HRs) for the association between average air pollution level and survival. Those studies all found strong associations between higher average air pollution levels during follow-up and decreased survival after a cancer diagnosis. In this case study, we will first use the approach from this previous work that creates a misalignment of the assigned exposure at time 0, where air pollution is averaged across follow-up and treated as time-fixed in a Cox model (naive approach). We then estimate this same association using the proposed discrete-time hazards approach (discrete-time approach). This is an established method for estimating the causal effect of a time-varying exposure in a time-to-event framework that upholds the principle of alignment at time 0 outlined by the target trial framework (18).

Study population

A retrospectively compiled, population-based cohort of lung cancer patients in California was created by linking data from the California Cancer Registry (CCR) with estimates of fine particulate matter air pollution (particulate matter less than or equal to 2.5 μm in aerodynamic diameter ($PM_{2.5}$)) using patient zip code at diagnosis. The CCR is a statewide population-based cancer surveillance system that collects information on all cancers (except nonmelanoma skin cancer) diagnosed in California. Lung cancer cases

diagnosed between 2000 and 2010 and registered by the CCR were identified with *International Classification of Diseases for Oncology, Third Edition* (ICDO-3) site codes (ICDO-3 codes C34.0–C34.3, C34.8, and C34.9). Cases were included in the study if the patient had a primary, histologically confirmed cancer of the lung or bronchus and had an address in California at the time of cancer diagnosis on record with the CCR (n = 123,706). Patients with in situ cancer (n = 48), with cancer diagnosed at autopsy (n = 355), or with an invalid date of diagnosis or follow-up (n = 2,035) were excluded from the analysis. For the purposes of this analysis, all cancer cases who survived for less than 30 days from the date of diagnosis were excluded (n = 15,625). The initial study population included 105,643 patients. This study was reviewed and approved by institutional review boards at San Diego State University and the University of California, San Diego, and by the California Department of Public Health Committee for the Protection of Human Subjects.

Exposure

Daily zip-code–specific concentrations of $PM_{2.5}$ ($\mu g/m^3$) were estimated from 24-hour daily mean values sampled and reported by the Environmental Protection Agency's Air Quality System in California. Measurements from the 3 nearest stations within 25 km of each population-weighted zip code centroid were assigned to each zip code week using inverse distance weighting as applied previously (19). The measured $PM_{2.5}$ concentration was weighted by the squared inverse distance to each zip code centroid, giving more weight to values at stations closer to the centroid. In this analysis, these estimated daily values were averaged as monthly zip code $PM_{2.5}$ values for each lung cancer patient. Patients with fewer than 2 daily values in all months of follow-up were excluded from the analysis (n = 16,605). In the discrete-time approach, which uses time-varying air pollution levels, patients are censored at their first month of follow-up with sparse air pollution data (less than 2 daily values). In a sensitivity analysis, the naive approach was replicated, additionally excluding individuals with any sparse air pollution months during follow-up (n = 4,099).

Outcome

Survival time was calculated from the date of lung cancer diagnosis to the date of death from any cause. All patients were censored 5 years after the date of diagnosis to avoid heterogeneity of effects associated with long-term survivorship (20).

Covariates

The following covariates were considered as potential confounders of the association between air pollution and lung cancer survival: age at diagnosis (years), sex (male/female), relationship status (single, partnered), race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian American, Native Hawaiian/Pacific Islander, other/unknown), treatment received during the first 6 months after diagnosis (surgery, radiation, chemotherapy), histological type, month of diagnosis, year of diagnosis, and socioeconomic status quintile. Socioeconomic status was measured using a composite residential neighborhood-level index that combined US Census measures of education, income, occupation, and cost of living at the census block group level, treated as quintiles (21). Standard histological groupings were created using ICDO-3 morphology codes for carcinoma, squamous cell carcinoma, adenocarcinoma, small cell carcinoma, large cell carcinoma, and other carcinomas.

Statistical analysis

Descriptive statistics were calculated for survival, $PM_{2.5}$ exposure, and other covariates. Two approaches were used to estimate the impact of air pollution on survival after a lung cancer diagnosis. The first approach (naive approach) emulates previous literature wherein $PM_{2.5}$ exposure is averaged over follow-up to create a single value that represents all exposure during the follow-up period. This average exposure is then treated as time-fixed in a Cox proportional hazards model to estimate a single HR for the effect of $PM_{2.5}$ on survival over the entire follow-up period.

The second approach (discrete-time approach) treats $PM_{2.5}$ as time-varying by using cumulative monthly averages. These monthly averages are then included in a discrete-time hazards model (i.e., a pooled logistic model with relatively short periods) where outcomes are assessed at each month of follow-up (22). This model estimates, in each month for each person, the conditional probability of remaining free of the outcome given cumulative average air pollution exposure, baseline covariates, and time of follow-up. In this approach, the discrete-time hazard during each month t is defined as the risk of the outcome during month t among participants who reached month t free of the outcome. Said differently, the analysis is stratified at each month during the follow-up period, and then month-specific effect estimates are pooled. This ensures that exposure is assigned at each outcome-time (month), which guarantees comparable exposure-time lengths for persons with and without the outcome at each month. This approach ensures alignment between treatment assignment and time 0 (for each case month distinct time 0 in the context of this time-varying discrete-time approach). Time-varying hazards were accounted for by modeling follow-up time using cubic splines. Bootstrap confidence intervals (CIs) were used to account for correlated observations.

All models for both approaches included potentially confounding variables that were associated with both $PM_{2.5}$ exposure and lung cancer survival, as described above, and included a random intercept on zip code at diagnosis to account for heterogeneity in survival across space. Models were stratified by disease stage at diagnosis (local, regional, or distant) due to established differences in etiology, treatment, and survival. Overall models included patients with an unknown stage of disease.

To illustrate the difference in effect estimates over follow-up between both methods, we estimated adjusted survival curves using terciles of $PM_{2.5}$ exposure. For the naive approach, we used the R software package "*survminer*"

Table 1. Baseline Characteristics of Lung Cancer Patients in the California Cancer Registry, by Disease Stage, 2000–2010

Characteristic	Disease Stage							
	Localized (*n* = 14,304)		Regional (*n* = 19,294)		Distant (*n* = 49,525)		Unknown (*n* = 5,915)	
	No.	%	No.	%	No.	%	No.	%
$PM_{2.5}$ concentration, μg/m^3[a]	12.6 (3.8)		12.9 (4.2)		13.3 (4.7)		14.1 (4.8)	
Age, years[a]	69.9 (11.3)		68.7 (11.0)		67.9 (11.7)		75.1 (11.7)	
Male sex	6,436	45.0	9,740	50.5	25,795	52.1	2,809	47.5
Histological type[b]								
Adenocarcinoma	6,904	49.3	7,376	38.8	19,252	40.0	961	25.7
Squamous cell carcinoma	3,040	21.7	4,679	24.6	6,486	13.5	579	15.5
Other carcinoma	2,438	17.4	3,356	17.7	9,569	19.9	631	16.9
Large cell carcinoma	962	6.9	1,401	7.4	5,259	10.9	1,140	30.4
Small cell carcinoma	670	4.8	2,183	11.5	7,615	15.8	433	11.6
Race/ethnicity[b]								
Non-Hispanic White	10,504	73.6	13,495	70.0	32,209	65.1	4,178	70.9
Asian/Pacific Islander	1,489	10.4	2,205	11.4	6,874	13.9	631	10.7
Hispanic	1,288	9.0	1,849	9.6	5,756	11.6	626	10.6
Non-Hispanic Black	952	6.7	1,641	8.5	4,473	9.0	436	7.4
American Indian	46	0.3	78	0.4	147	0.3	24	0.4
Partnered relationship status	7,696	54.8	10,455	55.3	26,013	53.9	2,306	40.7
Treatment								
Radiation	2,222	15.5	8,079	41.9	20,950	42.3	738	13.2
Surgery	9,649	67.5	8,160	42.3	2,407	4.9	180	3.2
Chemotherapy	1,833	12.9	9,405	49.7	26,298	54.2	954	18.1
SES quintile								
Low	1,910	13.4	2,891	15.0	8,113	16.4	1,003	17.0
Low medium	2,699	18.9	3,751	19.4	9,741	19.7	1,373	23.2
Medium	3,091	21.6	4,205	21.8	10,770	21.7	1,346	22.8
Medium high	3,199	22.4	4,242	22.0	10,878	22.0	1,226	20.7
High	3,405	23.8	4,205	21.8	10,023	20.2	967	16.3

Abbreviations; $PM_{2.5}$, particulate matter less than or equal to 2.5 μm in aerodynamic diameter; SES, socioeconomic status.
[a] Values are expressed as mean (standard deviation).
[b] Excludes persons with missing/unknown data.

(R Foundation for Statistical Computing, Vienna, Austria). In this package, a survival curve is plotted for each level of the grouping variable ($PM_{2.5}$ tercile) and the distribution is adjusted to the reference population using inverse probability weighting as described in the package documentation. For the discrete-time approach, we used the method described by Hernán (18). In this procedure, conditional survival is calculated by multiplying the model's predicted values through time t to estimate conditional survival at t for all subjects. Survival is then predicted at time t for each subject and conditional survival under each exposure is averaged for all subjects. This results in survival curves for each exposure level that are standardized to the distribution of the covariates in the study. All analyses were performed using R Studio, version 4.0.2.

RESULTS

The final study population included 89,038 lung cancer cases diagnosed in California between 2000 and 2010. Baseline characteristics of the study population are presented in Table 1. Lung cancer patients were, on average, 70.4 (standard deviation, 11.4) years of age at diagnosis and predominantly non-Hispanic White (68%). More than half of the lung cancers (55.6%) were diagnosed at a late disease stage.

During the 5-year study period, 82% of the population died, with the majority (53%) dying in the first year. Median durations of survival for localized, regional, and distant stage at diagnosis were 4.92, 1.69, and 0.53 years, respectively. Average $PM_{2.5}$ exposure was 12.4 μg/m^3 across all lung

Table 2. Estimated Hazard Ratios and Odds Ratios for 5-Year Cancer-Free Survival per 5-Unit (5-$\mu g/m^3$) Increase in $PM_{2.5}$ Exposure Over the Follow-Up Period, California Cancer Registry, 2000–2010

Approach	Unadjusted			Adjusted[a]		
	HR	Pooled OR	95% CI	HR	Pooled OR	95% CI
Naive						
Local	1.49		1.44, 1.55	1.37		1.31, 1.43
Regional	1.40		1.36, 1.43	1.32		1.28, 1.36
Distant	1.20		1.19, 1.22	1.18		1.16, 1.20
Overall	1.32		1.31, 1.34	1.38		1.36, 1.40
Discrete time[b]						
Local		1.05	1.02, 1.08		0.99	0.96, 1.02
Regional		1.04	1.02, 1.06		0.98	0.96, 1.00
Distant		1.03	1.02, 1.04		0.99	0.98, 1.01
Overall		1.03	1.03, 1.04		0.99	0.98, 1.00

Abbreviations: CI, confidence interval; HR, hazard ratio; OR, odds ratio; $PM_{2.5}$, particulate matter less than or equal to 2.5 μm in aerodynamic diameter.

[a] Adjusted for age at diagnosis (years), sex (male/female), relationship status (single, partnered), race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, Asian American, Native Hawaiian or Pacific Islander, other/unknown), treatment received during the first 6 months after diagnosis (surgery, radiation, chemotherapy), histological type, month of diagnosis, year of diagnosis, and socioeconomic status quintile. Overall estimates (all stages) additionally adjusted for disease stage at diagnosis (localized, regional, distant, unknown).

[b] ORs are for a 5-$\mu g/m^3$ increase in monthly $PM_{2.5}$ concentration over the course of follow-up. The 95% CIs were obtained via bootstrap estimation with 1,000 replications using the percentile method.

cancer cases. Over the study period, there were long-term downward trends in $PM_{2.5}$ concentration (Web Figure 3).

Table 2shows the estimated HRs for 5-year survival for both the naive approach and the discrete-time approach. In the naive approach, the HRs for all-cause mortality associated with a 5-$\mu g/m^3$ increase in average $PM_{2.5}$ exposure were 1.37 (95% CI: 1.31, 1.43), 1.32 (95% CI: 1.28, 1.36), and 1.18 (95% CI: 1.16, 1.20) for localized, regional, and distant disease stage, respectively. In the discrete-time approach, the odds ratios for all-cause mortality associated with a 5-$\mu g/m^3$ increase in time-varying $PM_{2.5}$ exposure were 0.99 (95% CI: 0.96, 1.02), 0.98 (95% CI: 0.96, 1.00), and 0.99 (95% CI: 0.98, 1.01) for localized, regional, and distant disease stage, respectively. Adjusted survival curves for both the naive and discrete-time approaches are shown in Figures 1 and 2. While there is substantial heterogeneity in the probability of survival by air pollution tercile in the naive curves, the discrete-time curves show relatively consistent survival probability by air pollution tercile.

Hazard ratios did not change substantially in a sensitivity analysis of the naive approach in which individuals with sparse air pollution months were excluded (Web Table 1).

DISCUSSION

While ITB is most often addressed in clinical epidemiology and pharmacoepidemiology, the errors that introduce systematic bias as a result of failure to correctly emulate a target trial can occur in environmental epidemiology as well. In this case study, we used the target trial framework to identify that averaging an environmental exposure over follow-up can introduce ITB into the study population and lead to substantial differences in the estimated effects and conclusions in the inference of interest. When an exposure is averaged over the course of follow-up, treatment assignment is made using postbaseline information, which unhitches treatment assignment from time 0 (7, 9). We illustrate that when ITB is introduced into a study in this way, the use of a time-to-event model does not adequately control for differences in follow-up lengths (a well-known function of these models). Instead, it is necessary to use alternative approaches that ensure alignment at time 0.

In this paper, we proposed an analytical solution using a discrete-time hazards approach. This method ensures alignment between time-varying treatment and time 0 at each outcome time because the hazard is reassessed in each follow-up period, including updating information on any time-varying covariates or exposures. This guarantees that only exposure periods that are comparable (with respect to time) are being used to assess the hazard at each outcome time. In a recent paper that outlined the use of causal diagrams for ITB, Mansournia et al. (23) highlighted how ITB can be considered either misclassification or selection bias and then proposed strategies for avoiding ITB, including the discrete-time approach we used here. They made the useful comparison of the discrete-time approach with a series of mini–randomized trials at monthly intervals (23). In this comparison, individuals have their exposure value reassigned in each mini–randomized trial (at each month), which prevents ITB.

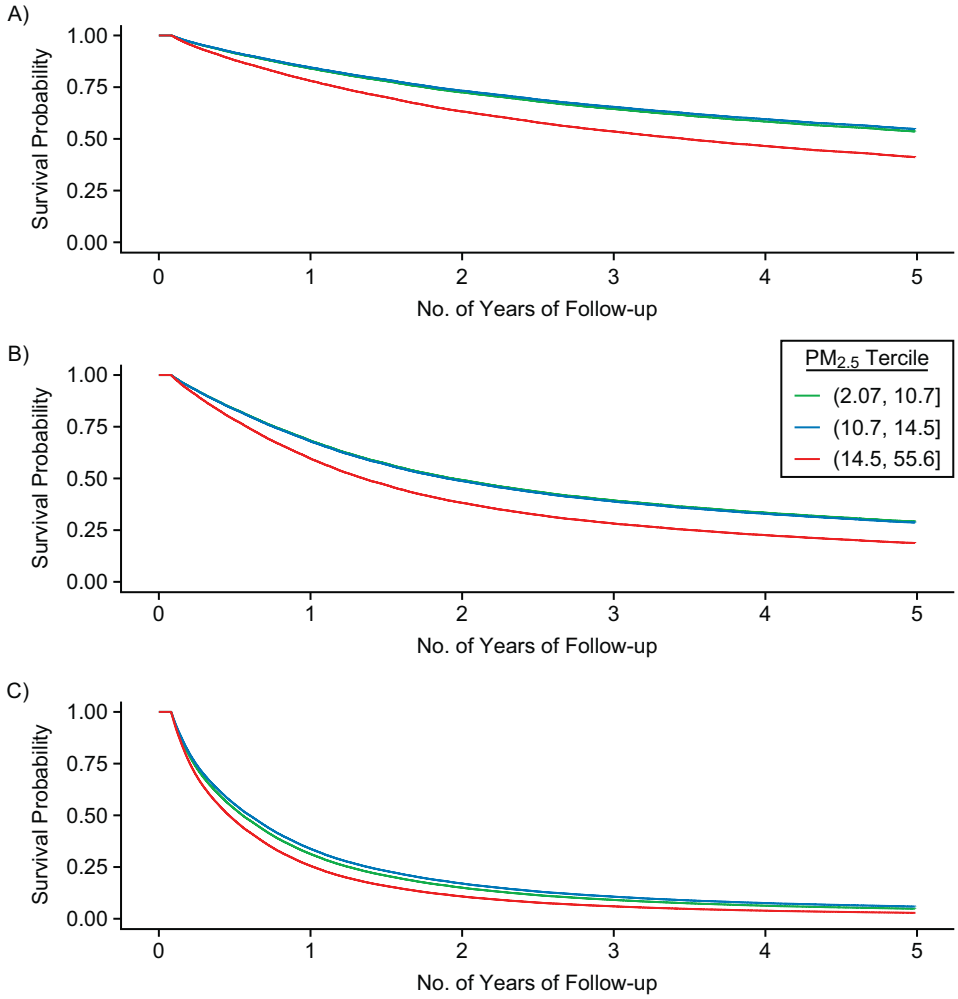

Figure 1. Adjusted survival curves from a naive analysis of the impact of air pollution on 5-year lung cancer survival, by disease stage, California Cancer Registry, 2000–2010. The analysis used average $PM_{2.5}$ exposure over the course of follow-up for localized (A), regional (B), and distant (C) stage of disease. Models were adjusted for age at diagnosis, sex, relationship status, race/ethnicity, treatment received during the first 6 months after diagnosis, histological type, month of diagnosis, year of diagnosis, and socioeconomic status quintile. Adjustment was done using inverse probability weighting in the "*survminer*" R software package. $PM_{2.5}$, particulate matter less than or equal to 2.5 μm in aerodynamic diameter.

Because discrete-time models allow for time-varying hazards and HRs, their use can also help investigators avoid common pitfalls in reporting causal effects using an average HR to represent risk over entire follow-up periods. Using an average HR can be uninformative because of time-varying HRs, which can occur as a result of built-in selection bias. The average HR ignores the distribution of events over follow-up, which guarantees that the magnitude of the HR will depend on the length of follow-up when the HR is time-varying. Hernán (18) recommends using discrete-time hazards models to obtain adjusted survival curves in order to present HRs over time, to overcome the misleading representation of a single HR that depends on length of follow-up. Furthermore, the discrete-time approach can be easily integrated into more advanced methods, including g-computation or inverse probability weighting methods (24, 25). These approaches could be particularly helpful when dealing with complex time-varying exposure and confounding settings that may cause exposure-confounding feedback (i.e., collider stratification bias) as described elsewhere (26, 27). While we chose to illustrate an analytical solution using discrete-time models for the aforementioned reasons, an alternative approach to avoiding ITB with environmental exposures has been demonstrated in previous studies, using standardized exposure windows (i.e., weeks, years) for all individuals, regardless of duration of follow-up (28, 29). Note that alternative analytical strategies could have been used to flexibly handle time-varying exposures and confounders in a discrete-time setting, such as doubly robust approaches including augmented inverse probability of treatment weighting or doubly matched estimators (30).

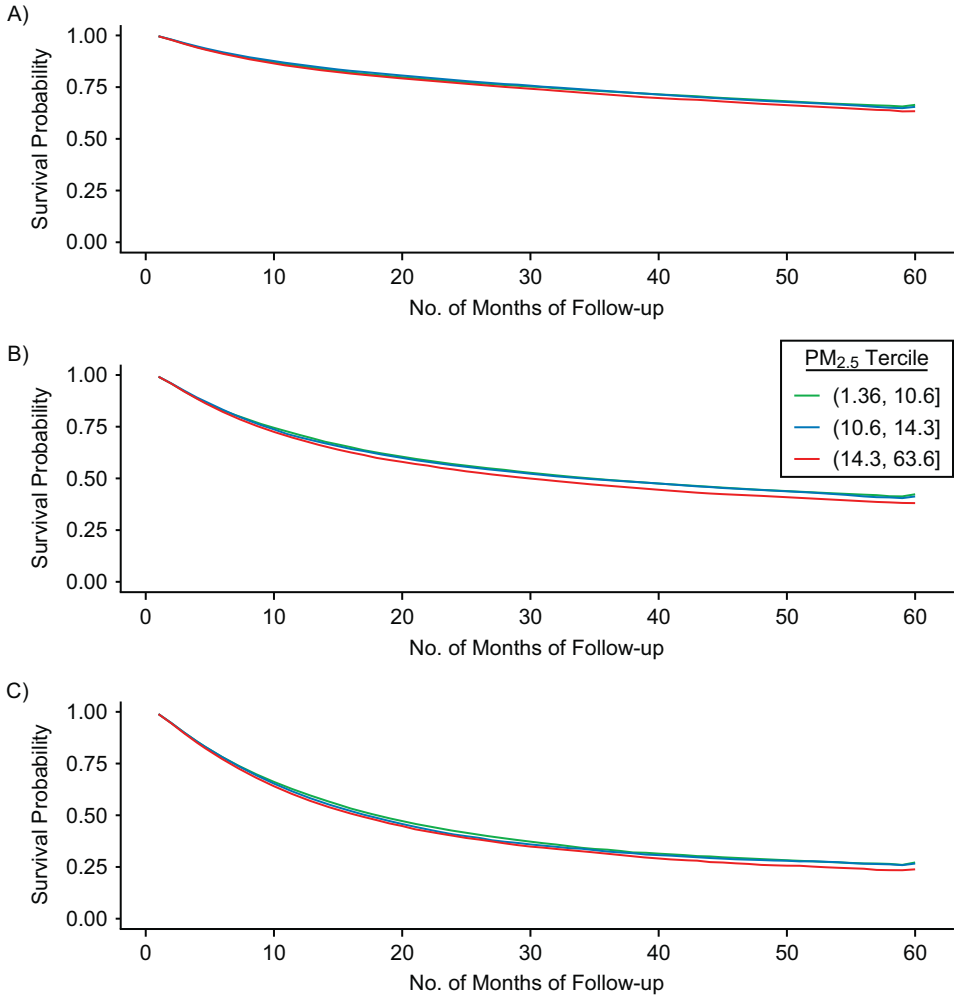

Figure 2. Adjusted survival curves from a discrete-time analysis of the impact of air pollution on 5-year lung cancer survival, by disease stage, California Cancer Registry, 2000–2010. The analysis used monthly $PM_{2.5}$ exposure over the course of follow-up for localized (A), regional (B), and distant (C) stage of disease. Models were adjusted for age at diagnosis, sex, relationship status, race/ethnicity, treatment received during the first 6 months after diagnosis, histological type, month of diagnosis, year of diagnosis, and socioeconomic status quintile. Adjustment was done using the approach described by Hernán (18). $PM_{2.5}$, particulate matter less than or equal to 2.5 μm in aerodynamic diameter.

After applying this discrete-time approach to our case study with lung cancer cases in California, we found little evidence that $PM_{2.5}$ exposure after a cancer diagnosis affects survival. By contrast, with the naive approach, which uses average exposure over follow-up, we found estimates that were similar in magnitude to results from the emulated studies (2, 4). The presence of ITB in the naive approach can be easily visualized by comparing the survival curves from the naive and discrete-time approaches. In a recent study examining the impacts of average air pollution exposure over follow-up on survival in California (5), investigators found that the HRs for all-cause mortality associated with a 1–standard-deviation increase in average $PM_{2.5}$ level were 1.38 (95% CI: 1.35, 1.41), 1.26 (95% CI: 1.24, 1.28), and 1.10 (95% CI: 1.09, 1.11) for localized, regional, and distant disease stage, respectively. We propose that these effect estimates are probably driven by ITB. Notably, the apparent effect modification by stage at diagnosis seen here and in other previous studies that used the naive approach (5) is no longer apparent after using the discrete-time approach. We hypothesize that this effect heterogeneity is likely to be driven by the larger differences in survival time among persons with a localized stage of disease, relative to those with a distant disease stage. These comparatively larger differences in follow-up time then exacerbate the impact of ITB.

In this case study, we demonstrated that ITB in this particular setting was substantial and was at least partially attributable to the strong downward trend in $PM_{2.5}$ over the study period (16, 17). Because of this temporal trend, the averaged exposure level was informed not only by the number of observations available (length of follow-up) but also by the decreasing trend itself. This means that individuals with longer follow-up times had systematically lower averages than those with shorter follow-up times, creating

an obvious difference in average exposure. Researchers in environmental epidemiology are aware of the strong temporal trends in most environmental exposures and often control for season and year in the analysis to adjust for this trend (2). Unfortunately, controlling for year of diagnosis using any strategy (stratification, indicator variable) will not control for temporal trends in the exposure or prevent ITB in the presence of an averaged exposure over follow-up. In this case study, the bias was particularly large because of the strong temporal trend in $PM_{2.5}$ concentration over the study period and the large differences between follow-up in lung cancer cases. While this bias may be minimized in other settings where the temporal trend is not as strong, it is still necessary to use approaches that ensure alignment at time 0 under the target trial framework to avoid systematic errors.

While we focused on a specific example in $PM_{2.5}$ exposure and lung cancer survival, ITB can occur in any observational setting. For example, in a recent paper that examined the impact of air pollution on mortality in a Medicare population, Di et al. (31) used yearly average air pollution exposure from the date of Medicare enrollment to the date of death or last follow-up in Cox proportional hazards models. Most recently, Tian et al. (32) examined the impact of air pollution and temperature on coronavirus disease 2019 (COVID-19) case fatality. The used average exposure from the date of diagnosis through the date of death or recovery. In both of these examples, air pollution exposure was assigned using attained length of follow-up. Because of this, these analyses may be vulnerable to ITB through the same mechanism as described in this case study. Because treatment assignment is not defined at time 0 inherently in observational studies, as it is in randomized trials, epidemiologists need to ensure this alignment explicitly to avoid these systematic errors.

There were several limitations in this work worth considering. First, we did not evaluate the conditions under which ITB would be introduced in environmental epidemiologic studies. ITB in environmental epidemiology may be subject to different considerations than its "classical" counterpart. Notably, ITB is most often identified in studies with a discrete exposure value, where subjects are considered either unexposed or exposed (targeting a traditional randomized trial with 2 arms and 1 single randomization at baseline). In air pollution and other environmental exposures, exposure is often measured continuously and is ubiquitous, which means no one is truly unexposed. In such settings, the target trial corresponds to multiple randomizations at different times, as in crossover or stepped wedge designs. There were additionally limitations in the case study itself. Air pollution exposure was assigned using interpolated estimates from ground-level fixed-site monitors, which are not evenly spaced across the state of California. This may have resulted in differential precision of exposure measurement by region, specifically in rural areas where there are fewer monitors. Because air pollution was assigned using residential zip code at diagnosis, we did not have information on individual variability in air pollution exposure from the participant's workplace or other factors that might influence daily exposure. The CCR does not include information on potentially important confounding variables, including smoking, alcohol use, access to health care (although insurance status can be an appropriate proxy), or treatment received more than 6 months after diagnosis.

In conclusion, in this study we identified that averaging a time-varying environmental exposure over follow-up in a time-to-event analysis creates misalignment between treatment assignment and the start of study follow-up, which can introduce ITB into the study. The bias resulting from this misalignment can be further exacerbated in the presence of time trends commonly found in environmental exposures such as air pollution. In this context, we recommend treating air pollution as time-varying in a model that ensures alignment between the start of study follow-up and treatment assignment. In future studies, researchers should evaluate under what conditions ITB in this setting is minimized or exacerbated.

ACKNOWLEDGMENTS

Author affiliations: Herbert Wertheim School of Public Health, University of California, San Diego, La Jolla, California, United States (Paige Sheridan); Climate, Atmospheric Sciences, and Physical Oceanography, Scripps Institution of Oceanography, La Jolla, California, United States (Chen Chen, Tarik Benmarhnia); Department of Epidemiology and Biostatistics, School of Public Health, College of Health and Human Services, San Diego State University, San Diego, California, United States (Caroline A. Thompson); and Palo Alto Medical Foundation Research Institute, Sutter Health, Palo Alto, California, United States (Caroline A. Thompson).

P.S. was supported by National Institutes of Health grants 1T32AG058529 and R01CA228147. The collection of cancer incidence data used in this study was supported by the California Department of Public Health, pursuant to the California Health and Safety Code (Section 103885); the Centers for Disease Control and Prevention's National Program of Cancer Registries, under cooperative agreement 5NU58DP006344; and the National Cancer Institute's Surveillance, Epidemiology and End Results Program, under contract HHSN261201800032I awarded to the University of California, San Francisco, contract HHSN261201800015I awarded to the University of Southern California, and contract HHSN261201800009I awarded to the Public Health Institute.

Information about the California Cancer Registry can be found at https://www.ccrcal.org/retrieve-data/data-for-researchers/. The California Cancer Registry shares data with researchers securely and confidentially for approved studies. Individual-patient–level data may be used to study cancer causes, prevention, treatment, and survival for research studies. Requests should be made by the principal investigator on behalf of the institution with which the investigator is affiliated. Environmental Protection Agency air quality data collected from ground site monitors can be found at https://www.epa.gov/outdoor-air-quality-data.

The ideas and opinions expressed herein are those of the author(s) and do not necessarily reflect the opinions of the

State of California, the California Department of Public Health, the National Cancer Institute, or the Centers for Disease Control and Prevention or their contractors and subcontractors.

Conflict of interest: none declared.

REFERENCES

1. Deng H, Eckel SP, Liu L, et al. Particulate matter air pollution and liver cancer survival. *Int J Cancer.* 2017;141(4):744–749.
2. Eckel SP, Cockburn M, Shu YH, et al. Air pollution affects lung cancer survival. *Thorax.* 2016;71(10):891–898.
3. Hu H, Dailey AB, Kan H, et al. The effect of atmospheric particulate matter on survival of breast cancer among US females. *Breast Cancer Res Treat.* 2013;139(1): 217–226.
4. Xu X, Ha S, Kan H, et al. Health effects of air pollution on length of respiratory cancer survival. *BMC Public Health.* 2013;13(1):800.
5. Villanueva C, Chang J, Ziogas A, et al. Ambient air pollution and ovarian cancer survival in California. *Gynecol Oncol.* 2021;163(1):155–161.
6. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016;183(8):758–764.
7. Hernán MA, Sauer BC, Hernández-Díaz S, et al. Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol.* 2016;79:70–75.
8. Gail MH. Does cardiac transplantation prolong life? A reassessment. *Ann Intern Med.* 1972;76(5):815–817.
9. Dickerman BA, García-Albéniz X, Logan RW, et al. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med.* 2019;25(10):1601–1606.
10. Adams RJ, Fuhlbrigge AL, Finkelstein JA, et al. Intranasal steroids and the risk of emergency department visits for asthma. *J Allergy Clin Immunol.* 2002;109(4):636–642.
11. Donahue JG, Weiss ST, Livingston JM, et al. Inhaled steroids and the risk of hospitalization for asthma. *JAMA.* 1997; 277(11):887–891.
12. Khurana V, Bejjanki HR, Caldito G, et al. Statins reduce the risk of lung cancer in humans: a large case-control study of US veterans. *Chest.* 2007;131(5):1282–1288.
13. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* 2007;16(3):241–249.
14. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol.* 2008;167(4):492–499.
15. Jones M, Fowler R. Immortal time bias in observational studies of time-to-event outcomes. *J Crit Care.* 2016;36: 195–199.
16. Bullock C, Ard K, Saalman G. Measuring the relationship between state environmental justice action and air pollution inequality, 1990–2009. *Rev Policy Res.* 2018;35(3): 466–490.
17. McClure CD, Jaffe DA. US particulate matter air quality improves except in wildfire-prone areas. *Proc Natl Acad Sci U S A.* 2018;115(31):7901–7906.
18. Hernán MA. The hazards of hazard ratios. *Epidemiology.* 2010;21(1):13–15.
19. Buteau S, Hatzopoulou M, Crouse DL, et al. Comparison of spatiotemporal prediction models of daily exposure of individuals to ambient nitrogen dioxide and ozone in Montreal, Canada. *Environ Res.* 2017;156:201–230.
20. Sugimura H, Yang P. Long-term survivorship in lung cancer: a review. *Chest.* 2006;129(4):1088–1097.
21. Yost K, Perkins C, Cohen R, et al. Socioeconomic status and breast cancer incidence in California for different race/ethnic groups. *Cancer Causes Control.* 2001;12(8):703–711.
22. Murray EJ, Caniglia EC, Petito LC. Causal survival analysis: a guide to estimating intention-to-treat and per-protocol effects from randomized clinical trials with non-adherence. *Res Methods Med Health Sci.* 2021;2(1):39–49.
23. Mansournia MA, Nazemipour M, Etminan M. Causal diagrams for immortal time bias. *Int J Epidemiol.* 2021;50(5): 1405–1409.
24. Breskin A, Edmonds A, Cole SR, et al. G-computation for policy-relevant effects of interventions on time-to-event outcomes. *Int J Epidemiol.* 2020;49(6):2021–2029.
25. Keil AP, Edwards JK, Richardson DB, et al. The parametric g-formula for time-to-event data: intuition and a worked example. *Epidemiology.* 2014;25(6):889–897.
26. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol.* 2016;46(2):756–762.
27. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550–560.
28. He J-R, Liu Y, Xia X-Y, et al. Ambient temperature and the risk of preterm birth in Guangzhou, China (2001–2011). *Environ Health Perspect.* 2016;124(7):1100–1106.
29. Bentayeb M, Wagner V, Stempfelet M, et al. Association between long-term exposure to air pollution and mortality in France: a 25-year follow-up study. *Environ Int.* 2015;85:5–14.
30. Leacy FP, Stuart EA. On the joint use of propensity and prognostic scores in estimation of the average treatment effect on the treated: a simulation study. *Stat Med.* 2014;33(20): 3488–3508.
31. Di Q, Wang Y, Zanobetti A, et al. Air pollution and mortality in the Medicare population. *N Engl J Med.* 2017;376(26): 2513–2522.
32. Tian F, Liu X, Chao Q, et al. Ambient air pollution and low temperature associated with case fatality of COVID-19: a nationwide retrospective cohort study in China. *Innovation (Camb).* 2021;2(3):100139.