Should We Adjust for Season in Time-Series Studies of the Short-Term Association Between Temperature and Mortality?

Lina Madaniyazi,[a] *Aurelio Tobías,*[a,b] *Ana M Vicedo-Cabrera,*[c,d] *Jouni J K Jaakkola,*[e,f] *Yasushi Honda,*[g,h] *Yuming Guo,*[i] *Joel Schwartz,*[j] *Antonella Zanobetti,*[j] *Michelle L Bell,*[k] *Ben Armstrong,*[l,m] *Michael J Campbell,*[n] *Klea Katsouyanni,*[o,p] *Andy Haines,*[l,m] *Kristie L Ebi,*[q] *Antonio Gasparrini,* [l,m,r] *and Masahiro Hashizume*[a,s]

Keywords: Mortality; Seasonal adjustment; Temperature; Time-series studies

Time-series studies of the short-term association between daily temperature and mortality are usually adjusted for seasonal confounding by functions of time in the regression analyses.[1–3] This adjustment aims to effectively separate seasonal patterns in the mortality series from the exposure–response association.[4–6] While such practice has become common, a recent Global Burden of Disease Study argued that "seasonal adjustments are not

[a]School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan;
[b]Institute of Environmental Assessment and Water Research, Spanish Council for Scientific Research, Barcelona, Spain;
[c]Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland;
[d]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland;
[e]Centre for Environmental and Respiratory Health Research, University of Oulu, Oulu, Finland;
[f]Finnish Meteorological Institute, Helsinki, Finland;
[g]Center for Climate Change Adaptation, National Institute for Environmental Studies, Tsukuba, Japan;
[h]Faculty of Health and Sport Sciences, University of Tsukuba, Tsukuba, Japan;
[i]Climate, Air Quality Research Unit, School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia;
[j]Department of Environmental Health, Harvard T H Chan School of Public Health, Harvard University, Boston, MA;
[k]School of Forestry and Environmental Studies, Yale University, New Haven, CT;
[l]Department of Public Health, Environments and Society, London School of Hygiene & Tropical Medicine, London, UK;
[m]The Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, UK;
[n]School of Health and Related Research, University of Sheffield, UK;
[o]National and Kapodistrian University of Athens, Medical School, Athens, Greece;
[p]Environmental Research Group, School of Public Health, Imperial College London, London, UK;
[q]Center for Health and the Global Environment, University of Washington, Seattle, WA;
[r]The Centre for Statistical Methodology, London School of Hygiene & Tropical Medicine, London, UK; and
[s]Department of Global Health Policy, School of International Health, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.

Lina Madaniyazi is an Assistant Professor at the School of Tropical Medicine and Global Health, Nagasaki University. Her research area includes assessing seasonality of health outcomes and the impact of environmental stressors on human health. The co-authors of this commentary have well-recognized expertise in the field of environmental epidemiology, contributing to the development of methodology for time-series regression and collaborating in relevant studies on the health effects of air pollution, temperature, and climate change for three decades. A.T., A.V.C., J.J., Y.H., Y.G., J.S., A.Z., M.B., B.A., K.K., A.G., and M.H. are members of the Multi-Country, Multi-City (MCC) Collaborative Research Network.

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

Data have been collected within the MCC (Multi-Country Multi-City) Collaborative Research Network (https://mccstudy.lshtm.ac.uk) under a data-sharing agreement and cannot be made publicly available. The R codes for the analysis are added in a GitHub repository, available at https://github.com/LinaMadaniyazi/Seasonal-adjustment.git

Correspondence: Lina Madaniyazi, School of Tropical Medicine and Global Health, Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan. E-mail: lina.madaniyazi@nagasaki-u.ac.jp.

ISSN: 1044-3983/23/343-313-318
DOI: 10.1097/EDE.0000000000001592

epidemiologically sound" and assessed the short-term risk of nonoptimum temperature on mortality without adjusting for seasonal confounding.[7,8]

In this commentary, we aim to provide an insight into the practice of seasonal adjustment in time-series regression models for the short-term relationship between temperature and mortality by illustrating statistical reasoning and discussing the underlying epidemiologic rationales.

Throughout, we will illustrate concepts and discussions through two examples based on real datasets. R code to reproduce our analyses are available on a GitHub repository (see Data Availability).

ILLUSTRATIVE EXAMPLES

We use data from two cities with distinct climate zones as examples. We collected daily counts of nonexternal mortality (International Classification of Diseases [ICD]-9 0-799, ICD-10 A00-R99) and daily mean temperature in two US cities, Detroit and Miami, in the periods 1985–2006. Detroit is a continental city with very cold winters and warm summers, while Miami is a tropical city with warm temperatures all year round. These datasets have been used in previous studies.[3,9] Here, temperature and mortality in both cities show a repeating seasonal pattern and a long-term trend (Figure 1), although the strength of these patterns is different between the two cities. Furthermore, Detroit shows a greater range in daily mean temperature than Miami (-24.4 to 31.4 in Detroit and 3.3 to 31.4 in Miami).

TIME-SERIES STUDIES IN ENVIRONMENTAL EPIDEMIOLOGY

The time-series regression is widely used to quantify short-term associations of exposures with outcomes in environmental epidemiology. It usually assumes a log-additive model to decompose the variation in the outcome variable at various timescales to isolate short-term variations that can be associated with a time-varying exposure. In the example datasets, we decomposed the time-series data on mortality into three components: long-term trend, seasonality, and random variation (eFigure 1; http://links.lww.com/EDE/C10). The long-term trend can be caused by demographic shifts or other slow changes. Seasonality, on the other hand, can be driven by a collection of factors that show seasonal variations, such as meteorology, influenza, Vitamin D levels, and sociobehavioral

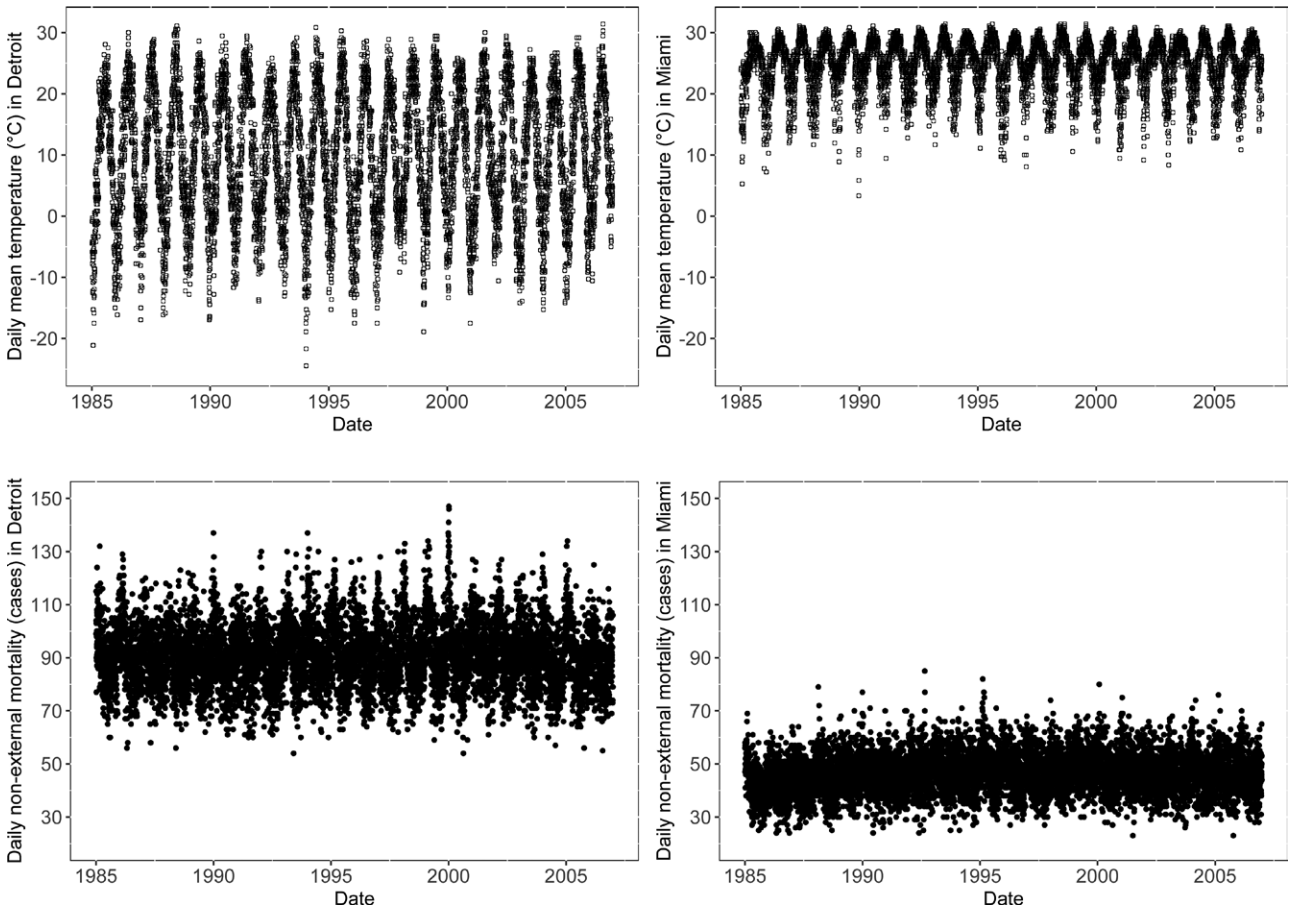

FIGURE 1. Daily time series of mean temperature (top) and nonexternal mortality (bottom) in Detroit (left) and Miami (right) from 1985 to 2006.

factors.[10] To guide our discussion, we use season to represent these seasonal factors in this paper, although season has many meanings in different settings.

As our interest in the time-series analysis is in short-term associations, the aim is to remove (i.e., control or adjust for) the first two components (i.e., long-term trend and seasonality) from the mortality series and examine whether temperature explains some of the remaining short-term variations in mortality. Otherwise, these two components can bias the short-term associations of interest and also lead to residual autocorrelation underestimating the standard errors of the parameters.[5] Thus, in a modern regression context, data analyst usually considers calendar time as a covariate and includes smooth functions of calendar time in models to remove long-term trend and seasonality, either separately by two functions or together by a single spline of time with enough flexibility to allow for both.[5]

By using our example datasets, we illustrated the short-term temperature–mortality association before and after controlling for the long-term trend and seasonality. We fitted a standard time-series Poisson regression model allowing for overdispersion. A cross-basis function was used to assess the short-term association between temperature and mortality.[6,11] Specifically, the cross-basis function is composed of two spline functions: one quadratic B-spline for the exposure–response association with three internal knots placed at the 10th, 75th, and 90th percentiles of location-specific temperature distribution, and one natural cubic spline for the lag–response association over lags days 0–21, with an intercept and three internal knots placed at equally spaced values in the log scale. We first assessed the temperature–mortality association without any adjustments. Then we included a natural cubic spline of date with 2 degrees of freedom (*df*) to control for long-term trends only and assessed the association. Finally, we updated the natural cubic spline of date with 8 *df* per year to control for long-term trend and seasonality simultaneously[3] and obtained a fully adjusted temperature–mortality association. We also examined the residual autocorrelation before and after the adjustments. Modeling choices were based on a previous study.[3]

The results suggested that the short-term temperature–mortality association can be affected by the adjustment of long-term trends and seasonality, especially the effect estimates of cold temperature (Figure 2; eTable 1 http://links.lww.com/EDE/C10). However, this impact may depend on study locations and the length of the study period. Notably, the confidence intervals are narrower before adjustments. Also, there is initially large residual autocorrelation which is minimized after the adjustment of both long-term trend and seasonality (eFigure 2; http://links.lww.com/EDE/C10). In addition, the quasi-Akaike information criterion favored the adjusted model (eTable 1; http://links.lww.com/EDE/C10).

CAUSAL STRUCTURES BETWEEN SEASON AND TEMPERATURE

The statistical rationale for removing seasonality in time-series regression studies is well accepted, whereas epidemiologic reasoning has not been explicitly stated. As noted

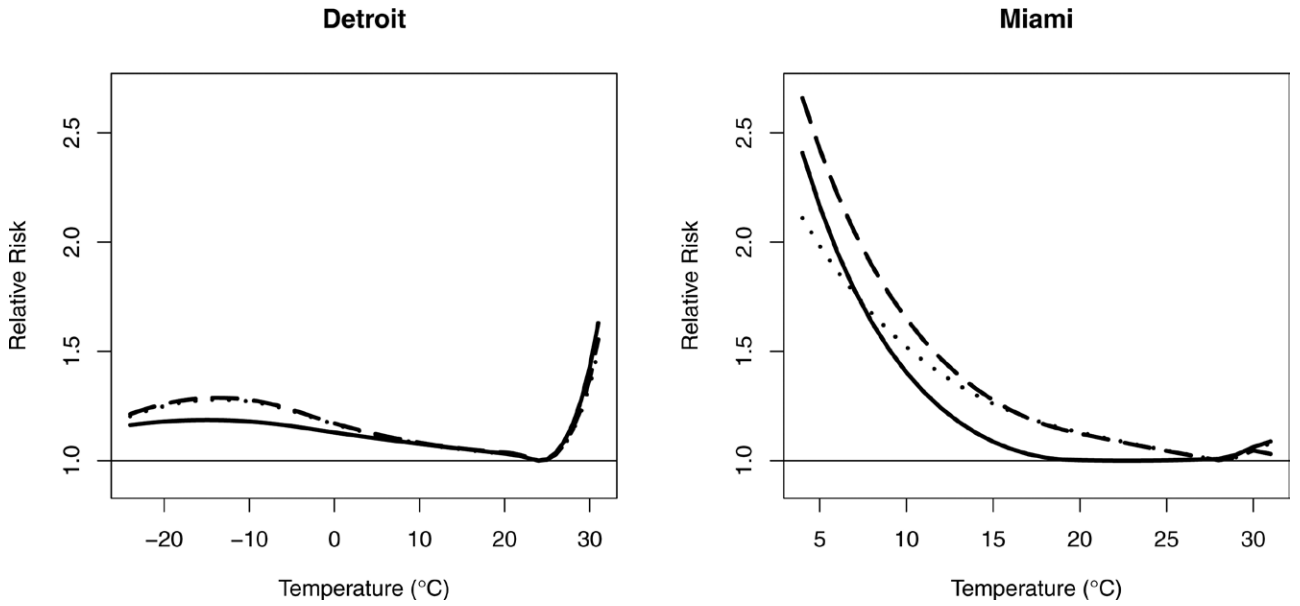

FIGURE 2. The association between temperature and nonexternal mortality before and after the adjustment of the long-term trend and seasonality in Detroit and Miami. Short-dashed line: temperature–mortality association without any adjustments. Long-dashed line: temperature–mortality association with adjustment of long-term trend by using a natural cubic spline of date with 2 degrees of freedom (*df*). Solid line: temperature–mortality association with adjustment of long-term trend and seasonality simultaneously by using a natural cubic spline of date with 8 *df* per year.

above, seasonality is one of the time components in time-series data, driven by a collection of seasonal factors. Therefore, the removal of seasonality in time-series regressions can be interpreted as the adjustments of seasonal effects in epidemiology. Here, we use directed acyclic graphs (DAG)[12,13] to discuss three potential causal structures for the relationship among temperature (T), season (S), and mortality (M). To reiterate, season here is used to represent the collection of seasonal factors driving the seasonality of mortality. It should be noted that the DAGs in Figure 3 are simplified depictions of potentially complicated causal relationships between temperature, season, and mortality. The relationships of the season with temperature may be very complex: some factors and temperature may affect each other (e.g., humidity), while some might be affected by temperature (e.g., influenza).

The rationale for the adjustment of season has been described as a concern about potential confounding.[4,6] Figure 3A illustrates a causal structure where season is a confounder of the temperature–mortality association. Then it would be necessary to adjust for season to obtain an unconfounded estimate of the temperature–mortality association. Figure 3B, on the other hand, presents a causal structure where temperature affects both season and mortality. Here, season is on an intermediate pathway between temperature and mortality. Thus, season is not a confounder of the temperature–mortality association since season does not affect temperature. In this case, adjusting for the season would give the direct effect of temperature unmediated by season, and the logic behind it is questionable if the interest is to assess the total effect of temperature. Figure 3C illustrates a causal structure where season lies upstream of the causal pathway from temperature and mortality, and all the effects of the season are mediated through temperature. There would be no need for seasonal adjustment; otherwise, it may diminish the statistical power and reduce precision in the estimates of the temperature effect.[14]

What causal structures between temperature and season would justify the adjustment for the season in time-series studies of the short-term effect of temperature then? The answer to this question should be dictated by hypotheses on how the season is causally related to temperature and mortality. As noted above, the season is a proxy for unmeasured variables that show seasonal patterns (e.g., infectious diseases, air pollution, and changes in dietary and behavioral patterns) and that can be risk factors on mortality; however, their mechanism as risk factors are not solely via temperature. Therefore, Figure 3C is not likely plausible. How do we choose between Figures 3A and 3B? Temperature and season are correlated: some seasonal variables, such as the changes in tropospheric ozone concentrations, influenza activity, and behavioral patterns, might be impacted by temperature, whereas some others (e.g., intensity and duration of sunlight) may impact temperature in turn. The data can also shed light on this. We assessed seasonality in mortality in Detroit and Miami before and after the adjustment of temperature over lag days 0–21 described above, while the seasonality of mortality was assessed using a cyclic spline function with 4 *df*.[15] Here, we observed a reduction though not elimination in seasonality after the adjustment (Figure 4), suggesting that temperature can be a partial mediator (Figure 3B) or a confounder (Figure 3A) for the season-mortality association. It also suggests that Figure 3C is not plausible since the seasonality in mortality did not eliminate after temperature adjustment. Similar results were

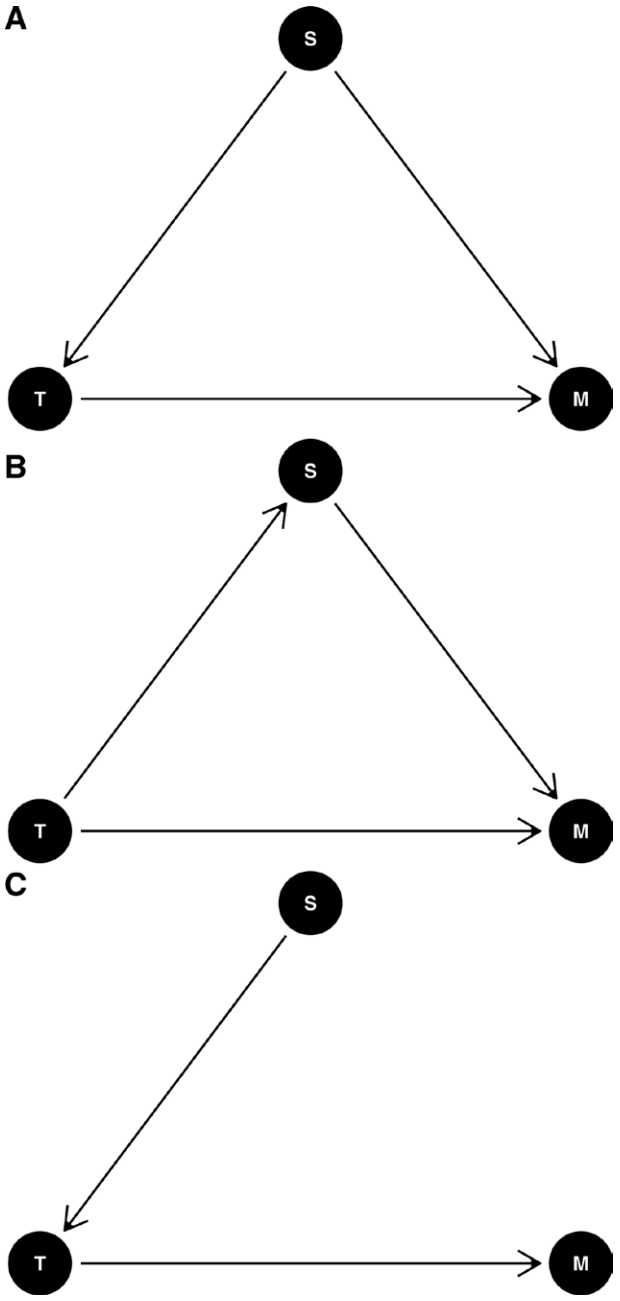

FIGURE 3. Directed acyclic graphs for the relationship between temperature (T), season[a] (S), and mortality (M). [a]Season is a surrogate or placeholder for any number of factors that vary seasonally.

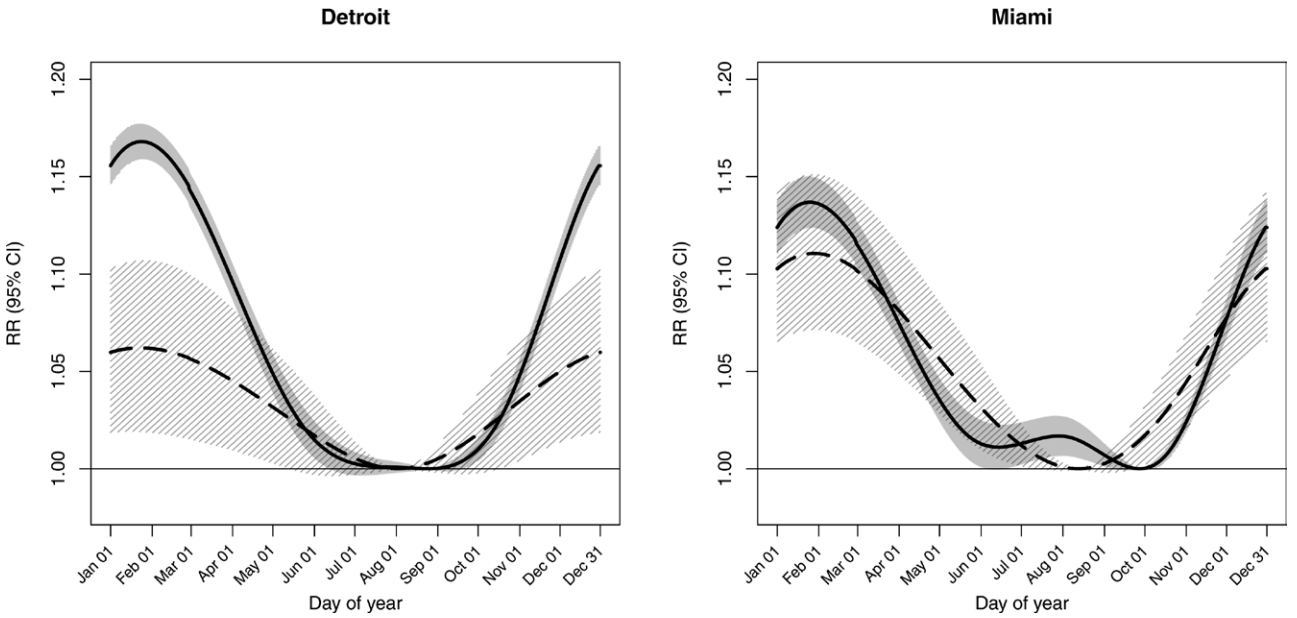

FIGURE 4. Seasonal variation in nonexternal mortality before (solid) and after (dashed) removing temperature effect.

reported in our previous multicountry multicity study.[9] Thus, Figures 3A,B are plausible.

In time-series regression studies of the short-term association between temperature and mortality, it is of interest to adjust for those seasonal confounders. However, these seasonal confounders are usually difficult to identify and measure. Thus, the data analyst adjusts the confounding effect of the season by using a function of calendar time (e.g., day-of-year, week, and month).[5,6,16] In other words, such a function is used to represent those seasonal variables that are on the open back-door pathway between temperature and mortality (Figure 3A). Such a causal model is justifiable. For a variable to be a confounder, it needs not to cause the exposure to introduce bias.[17] One firm requirement, however, is that it must not be on the causal pathway between exposure and outcome. Time can predict but cannot be caused by the variation in daily mean temperature or daily mortality. Thus, a function of time cannot be on the causal pathway as a mediator. In that case, the practice of using a function of time to adjust the confounding effect of the season is reasonable, and the absence of such an adjustment is questionable.

In our example, our assessment in Detroit and Miami shows that the adjustment of seasonal confounding by a function to time reduced estimates of cold effect but increased estimates of heat effect in both cities (Figure 1 and eTable 1; http://links.lww.com/EDE/C10). Thus, ignoring seasonal confounding in time-series analysis of the short-term effect of temperature may overestimate the cold effect while underestimating the heat effect. Such a bias may be more substantial when the exposure of interest is extreme temperature events, because, by not conditioning on seasonal time scales, we are comparing deaths on event days to deaths on all other days, including in the other seasons.

CONCLUSION

This commentary illustrates the importance of epidemiologic reasoning in the adjustment of seasonal confounding in time-series studies for the short-term relationship between temperature and mortality. However, the practice of the adjustment depends on the specific exposure and outcome, the study design being applied, and the covariates being measured. In case–crossover studies, for example, the adjustment of seasonal confounding can be omitted, as it is controlled for by design.[18] Further thinking about seasonal adjustment is merited in panel studies of temperature effects where health outcomes are usually measured for a short time period but across different seasons.

Adjustment for seasonal confounding typically is warranted when estimating unconfounded effect estimates of temperature on mortality. Otherwise, investigators should clarify their assumptions on causal relationships for omitting seasonal confounding in their analysis. This will help the readers to interpret the effect estimates properly.

REFERENCES

1 Campbell MJ. Time series regression for counts: an investigation into the relationship between sudden infant death syndrome and environmental temperature. *J R Stat Soc Ser A Statistics Soc*. 1994;157:191.
2 Anderson BG, Bell ML. Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*. 2009;20:205–213.
3 Gasparrini A, Guo Y, Hashizume M, et al. Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *Lancet*. 2015;386:369–375.

4 Gasparrini A, Armstrong B. Time series analysis on the health effects of temperature: advancements and limitations. *Environ Res*. 2010;110:633–638.
5 Bhaskaran K, Gasparrini A, Hajat S, Smeeth L, Armstrong B. Time series regression studies in environmental epidemiology. *Int J Epidemiol*. 2013;42:1187–1195.
6 Armstrong B. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*. 2006;17:624–631.
7 Burkart KG, Brauer M, Aravkin AY, et al. Estimating the cause-specific relative risks of non-optimal temperature on daily mortality: a two-part modelling approach applied to the Global Burden of Disease Study. *Lancet*. 2021;398:685–697.
8 Burkart KG, Brauer M, Aravkin AY, et al. Global mortality burden attributable to non-optimal temperatures – Authors' reply. *Lancet*. 2022;399:1113–1114.
9 Madaniyazi L, Armstrong B, Chung Y, et al. Seasonal variation in mortality and the role of temperature: A multi-country multi-city study. *Int J Epidemiol*. 2022;51:122–133.
10 Peng RD, Dominici F. *Statistical methods for environmental epidemiology with R: a case study in air pollution and health*. 2008: 144.
11 Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. *Stat Med*. 2010;29:2224–2234.
12 Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995;82:702669–702710.
13 Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*1999;10:37–48.
14 Schisterman EF, Cole SR, Platf RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology*. 2009;20:488.
15 Madaniyazi L, Tobias A, Kim Y, Chung Y, Armstrong B, Hashizume M. Assessing seasonality and the role of its potential drivers in environmental epidemiology: a tutorial. *Int J Epidemiol*. 2022;51:1677–1686.
16 Imai C, Armstrong B, Chalabi Z, Mangtani P, Hashizume M. Time series regression model for infectious disease and weather. *Environ Res*. 2015;142:319–327.
17 Rothaman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Lippincott Williams & Wilkins; 2008.
18 Barnett AG, Dobson AJ. *Analysing Seasonal Health Data*. Springer; 2010.