Title: Evaluating the suitability of the case-crossover design under changing baseline outcome risk: A simulation of ambient temperature and preterm birth

Abstract: Case-crossover (CCO) studies are case-only within-person analyses of associations between acute exposures and outcomes. By design, the CCO is not confounded by time-invariant characteristics. CCO studies assume stable baseline outcome risks. Since the baseline risk of birth increases secularly over gestation, the validity of CCOs of preterm birth (PTB) is not clear. We simulated associations between temperature and PTB in New York State using historical ambient temperature data for LaGuardia Airport and PTB data from the CDC using different control period durations. CCO analyses were conducted with conditional logistic regression. We calculated bias according to the absolute difference between the simulated and estimated effects in the natural log scale and found that 1-month stratified control period selection yielded bias away from the null across all simulated effects (median = 0.018, IQR: 0.010 - 0.026). In contrast, the 2-week stratification resulted in negligible bias (median: 0.001, IQR: -0.011 - 0.012). Coverage of the 95% confidence intervals decreased with higher effects for the 1-month stratification, with a range 17-87% of estimated intervals including the simulated effect. 2-week stratification had consistent coverage across models (range 95-96%). Our findings suggest that future studies should consider using shorter time strata for PTB.

Authors: Daniel Carrión*[1], Johnathan Rush[1], Elena Colicino[1,2], Allan C. Just[1,2]

Affiliations:
*Corresponding author
[1] Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY
[2] Institute for Exposomic Research, Icahn School of Medicine at Mount Sinai, New York, NY

Keywords: case-crossover; preterm birth; simulation; extreme temperature; changing baseline risk; time stratified

Background

The case-crossover (CCO) design is widely used in epidemiology. CCO studies are case-only, within-person comparisons, therefore not confounded by time-invariant characteristics.[1] Proper statistical inference with CCO studies relies on appropriate selection of control time periods, and the time-stratified design is a preferred method for control selection.[2] However, CCO studies assume stable baseline outcome risks.[3] Multiple studies have used the CCO approach linking ambient temperature and preterm birth (PTB),[4] but since the baseline PTB risk increases secularly over gestation, the validity of the CCO design for PTB is not clear.[5] We utilized simulation methods to assess the appropriateness of the CCO in associations between ambient daily temperature and PTB.

Methods

We conducted a simulation using 2018 data for New York State. Data were acquired from National Weather Service records for LaGuardia Airport (daily maximum temperature) and the Centers for Disease Control's WONDER database (PTB data).[6] Daily births by gestational age are unavailable so were estimated using marginal distributions of births by month and day of week (within month).

Estimated daily births per gestational age (20-36 weeks) served as the basis for baseline risk. Baseline risks were combined with simulated (true) temperature effects to create expected counts per day. Relative risks (RR) were set from 0.9 to 1.25 per 10°F increase on lag day 0. Repeated Poisson random number generation created 1000 datasets per simulated RR.[7] Counts were disaggregated into individual records for CCO analyses, estimated via conditional logistic regression with 2-week and 1-month stratified control period selection matched on the day of week. Analyses only included warm months, May through September.[4] Sensitivity analyses included: 1) modeling 2007 data and 2) adjusting for gestational age or the proportion of month.

We summarized results according to bias and coverage. We calculated bias according to the absolute difference $ln(\hat{\beta}) - ln(\beta)$, where $\hat{\beta}$ is the estimated effect and β is the true effect. Coverage is the proportion of simulations where the 95% confidence intervals include the true effect. Reproducible analyses were conducted in R 4.0.2.[8]

Results

The 2-week stratified model exhibited little bias (**Figure 1A**), with a median bias of 0.001 (IQR: -0.011 - 0.012) across all simulated effects. Bias away from the null (median 0.018, IQR: 0.010 - 0.026) was identified for 1-month-stratified models. For example, the median odds ratios for a simulated relative risk of 1.05 per 10F were 1.05 for the 2-week stratified model and 1.07 for the 1-month stratified model.

Coverage of 95% confidence intervals was lowest for month-stratified results; between 17-87% of all intervals included the simulated effect (**Figure 1B**). The coverage was highest for 2-week stratified models, with ranges of 95-96%.

Sensitivity analyses simulating data from 2007 were consistent in bias and coverage (Supplemental materials). Adjustments for gestational age or proportion of month did not improve bias in 1-month stratified analyses.

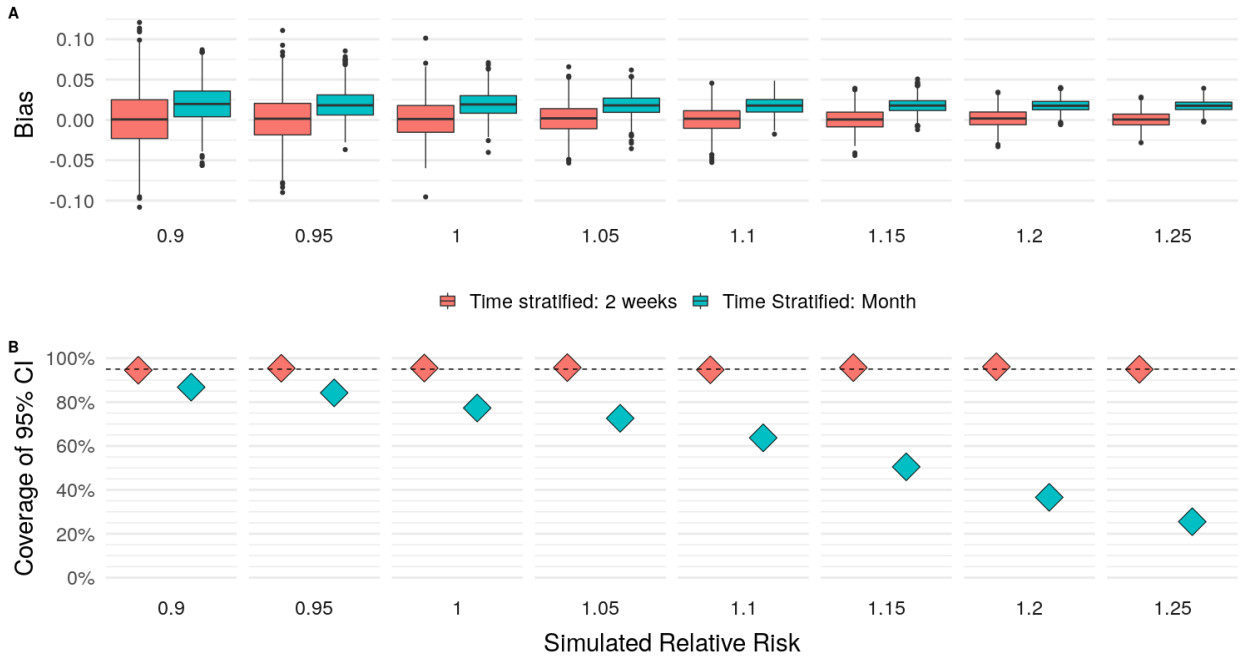

Figure 1: Results from 2018 simulations and case-crossover analyses. Colors represent analysis type, namely 2-week stratified and 1-month stratified control selection. **A**) Distribution of bias for each simulated effect, scaled to a 10F increase in temperature. **B**) The proportion of simulations where confidence intervals contain the simulated effect. Dashed line at 95%.

Conclusions

CCO studies may suffer systematic bias if selected control times are non-exchangeable within person.[1] While temperature varies seasonally and the risk of PTB changes quickly over pregnancy, we show that the use of fixed 2-week strata results in negligible bias and appropriate coverage of confidence intervals. Month-stratified models perform poorly, demonstrating consistent bias away from the null and decreasing coverage of 95% confidence intervals with larger effects. Many studies in this field have used month stratification, thus may be vulnerable to these biases. Future work should use shorter control windows for proper inference.

Data availability: All data and analytical code to reproduce results are available on GitHub: https://github.com/justlab/casecrossover_preterm_simulation.

Funding: This work was supported by NIH grant P30ES023515. DC is funded by NIH T32HD049311.

References

1. Mittleman MA, Mostofsky E. Exchangeability in the case-crossover design. *Int J Epidemiol*. 2014;43(5):1645-1655. doi:10.1093/ije/dyu081
2. Carracedo-Martínez E, Taracido M, Tobias A, Saez M, Figueiras A. Case-Crossover Analysis of Air Pollution Health Effects: A Systematic Review of Methodology and Application. *Environ Health Perspect*. 2010;118(8):1173-1182. doi:10.1289/ehp.0901485
3. Maclure M. The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. *Am J Epidemiol*. 1991;133(2):144-153. doi:10.1093/oxfordjournals.aje.a115853
4. Chersich MF, Pham MD, Areal A, et al. Associations between high temperatures in pregnancy and risk of preterm birth, low birth weight, and stillbirths: systematic review and meta-analysis. *BMJ*. Published online November 4, 2020:m3811. doi:10.1136/bmj.m3811
5. Darrow LA. Invited Commentary: Application of Case-Crossover Methods to Investigate Triggers of Preterm Birth. *Am J Epidemiol*. 2010;172(10):1118-1120. doi:10.1093/aje/kwq327
6. Centers for Disease Control & Prevention. CDC WONDER. Published 2021. Accessed November 8, 2019. http://wonder.cdc.gov
7. Lu Y, Zeger SL. On the equivalence of case-crossover and time series methods in environmental epidemiology. *Biostatistics*. 2007;8(2):337-344. doi:10.1093/biostatistics/kxl013
8. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2020. https://www.R-project.org/