

## When the Answer is "Big(ger) Data in Environmental Epidemiology: What are the Questions?"

Kaufman J <sup>1</sup>, Brauer M <sup>2</sup>, Samoli E <sup>3</sup>

<sup>1</sup> University of Washington, <sup>2</sup> University British Columbia, <sup>3</sup> University of Athens

Definitive answers to several types of Environmental Epidemiology research questions appear to require the use of ever larger data sets. In addition to meta-analytic approaches, researchers are increasingly turning to pooling smaller cohorts and to using administrative datasets. While these approaches offer increased statistical power, which can overcome some challenges of individual cohort studies, they introduce other methodological challenges and the potential for bias. This symposium will review these methodological challenges and the approaches that can be used to mitigate the potential for important biases.

- 1) Overall analytic approaches (pooling individual data vs meta-analysis/meta-regression)
- 2) Missing or differential data on potential confounders
- 3) Suboptimal data on health outcomes
- 4) Non-random attrition in cohorts
- 5) Non-closed cohort concerns in administrative records
- 6) Precision of exposure assignment in large populations
- 7) Privacy issues
- 8) Differential time periods for cohort enrollment
- 9) Differences in quality of confounder measurement

Presenters will review standard and novel approaches to these methodological challenges. The symposium will include discussion of advantages and disadvantages of solutions.

The future of environmental epidemiology requires adjusting to new realities of research infrastructure. One key issue, understanding health effects at low-levels, requires large sample sizes to understand effects--which increasingly will require taking advantage of research platforms created for other purposes. Small effects, high risk of residual confounding, outcome classification, exposure measurement error. There are many threats to valid results--the same risks that plagued the history of Environmental Epidemiology are present in our future.