Bayesian Measurement Error Models for air pollution and health studies

Erica Ponzi[1,2], Paolo Vineis[2], and Marta Blangiardo[2]

[1]Department of Biostatistics, Epidemiology, Biostatistics and Prevention Institute, University of Zürich, Hirschengraben 84, 8001 Zürich, Switzerland
[2]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom

Abstract

Studies in air pollution epidemiology require measurements of pollutants and more generally of environmental exposures, which may be affected by low precision and measurement error. This is a possible cause of bias in the estimation of parameters relevant to the study and can lead to inaccurate conclusions when evaluating associations among pollutants, disease risk and biomarkers. Although the presence of measurement error in such studies has been recognized as a potential problem, it is rarely considered in applications and practical solutions are still lacking. In this work, we formulate Bayesian measurement error models and apply them to study the link between air pollution and chronic cardiovascular diseases, focusing on how this link is mediated by omics measurements. The data stem from the "Oxford Street Study", a randomized crossover trial in which 60 volunteers walked for two hours in a traffic-free area (Hyde Park) and in a busy shopping street (Oxford Street) of London. Omics measurements are taken on each individual as well as air pollution measurements, in order to investigate the association between short-term exposure to traffic related air pollution and perturbation of metabolic pathways. Pollutants measures are likely to be affected by measurement error, which potentially propagates to the parameter estimates and may lead to biased estimations and conclusions concerning the casual relationship among omics data, environmental variables and disease risk. We implemented error-corrected models in a classical and a Berkson measurement error framework. Models are implemented using traditional `MCMC` simulative methods as well as `INLA` and their performance is compared in terms of accuracy and computation time.