Bayesian profile regression mixture models to estimate an instantaneous excess risk of cancer from highly correlated exposures and censored survival data.

Sophie Ancelet[1], Marion Belloni[1], Olivier Laurent[1], Chantal Guihenneuc[2]

[1] Institut de Radioprotection et de Sûreté Nucléaire, PSE-SANTE/SESANE/LEPID, Paris, France

[2] Université de Paris, Unité de Recherche “Biostatistique, Traitement et Modélisation des données biologiques” BioSTM - UR 7537, Paris, France

E-mail for correspondence: sophie.ancelet@irsn.fr

Abstract: In this work, we focused on the problem of estimating a disease risk from a few highly correlated environmental exposures and a highly censored survival outcome. We extended Bayesian profile regression mixture (PRM) models to this context by assuming an instantaneous excess hazard ratio disease sub-model and conducted a simulation study to explore the performance of these models to deal with co-exposures. Our hierarchical model incorporates a truncated Dirichlet process mixture as an attribution sub-model. An adaptive Metropolis-Within-Gibbs algorithm, including label-switching moves, was implemented to infer the model. This allows simultaneously clustering individuals with similar risks and similar exposure characteristics and estimating the associated instantaneous excess risk for each group. Our Bayesian PRM model was applied to the estimation of the risk of death by lung cancer in a cohort of French uranium miners who were chronically and occupationally exposed to multiple and correlated sources of ionizing radiation. This case study shows that PRM models are promising tools for exposome research and opens new avenues for methodological research in this class of models. It also highlights the potential limit of using standard MCMC algorithms to fit these models, even if the updating schemes for the class labels incorporate label-switching moves.

Key words: Bayesian inference, ionizing radiation, lung cancer, multicollinearity, truncated Dirichlet process mixture

Wild CP. (2005) Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomark Prev.. 14:1847–50. doi: 10.1158/1055-9965.EPI-05-0456

Rage E, Caër-Lorho S, Drubay D, Ancelet S, Laroche P, Laurier D. (2015) Mortality analyses in the updated French cohort of uranium miners (1946–2007). Int Archiv Occupat Environ Health. (2015) 88:717–30. doi: 10.1007/s00420-014-0998-6

Molitor J, PapathomasM, JerrettM, Richardson S. (2010) Bayesian profile regression with an application to the National Survey of Children’s Health. Biostatistics. 11:484–98. doi: 10.1093/biostatistics/kxq013