Taking into Account Spatial Autocorrelation in ecological study. An environmental health illustration: Proximity to Industries, Neighborhood Deprivation and Infant Mortality

Benoit Lalloué[1,2], Cindy Padilla[1,2], Denis Zmirou-Navier[1,2,3,4], Deguen Séverine[1,2]

[1] Ecole des Hautes Etudes en Santé Publique – Rennes- Sorbonne Paris Cité - France
[2] Institut de Recherche en Santé Environnement Travail-France
[3] INSERM U954–Vandoeuvre-les-Nancy-France
[4] Université de Nancy –Vandoeuvre-les-Nancy-France

Introduction: Ecological studies seldom include spatial auto-correlation of the statistical units, although not taking it into account can lead to biased estimations in statistical models.

Methods: One public health objective of the present study is to explore the role of the neighbourhood socio-economic status in the potential association of infant mortality rate with proximity to polluting industries, in the Lille metropolitan Area, between 2000 and 2009. The geographical unit used is the smallest census tract: IRIS. A deprivation index was calculated across the IRIS, from the INSEE census data of 1999. We use the Gaussian auto-regressive hierarchical "BYM" model to take into account both the high variability in mortality rates in each IRIS and the spatial autocorrelation between the statistical units. Different models, with and without spatial components, and with and without covariates, were used to compare the influence of spatial autocorrelation on the resultant statistical associations. Deviance Information Criterion (DIC) was used to compare models' goodness-of-fit (the lower the DIC, the better the model).

Results: The results show that the risk of infant mortality is multiplied by 1.53 (95% CI [1.33 to 1.76]) for a unit increase of deprivation and no significant association with proximity to polluting industries whatever the model. Moreover, the mortality risk increases with the increase of the deprivation level with and without spatial components taken into account, but relative risks are generally overestimated in the model without spatial components. Finally, models taking into account spatial components improve the models' goodness of fit in comparison with models without spatial components (DIC= 1509.690 vs 1580.180 without covariates; DIC= 1498.310 vs 1519.090 with covariates)

Conclusion: Taking into account spatial autocorrelation is very important in ecological study. Ignoring the spatial autocorrelation in this kind of models may produce biased and/or uncertain estimates and lead to wrong conclusions of significant associations where there are not and reciprocally. This study is another warning to researchers using ecologic studies in public health field.

Keywords: *Health Inequalities- Infant mortality- Deprivation index – Small areas – Census tract – Pollution – Spatial autocorrelation*