

Invited session - Epidemiology

Data integrating in viral evolutionary inference: applications to spatial genetics problems

Philippe Lemey (Rega Institute, K.U. Leuven, e-mail: philippe.lemey@rega.kuleuven.be)

Summary: Combating pathogen spread and the associated disease burden is a tremendous challenge requiring sustained research effort and decided public health measures, and the availability of genomic data provides a major asset in characterizing these pathogens. Recent developments in pathogen phylodynamics aim at a marriage of statistical thinking and evolutionary biology to integrate these data through phylogenetic reconstructions with host, phenotypic and geographic sampling information.

Here, I will highlight recent advances in Bayesian evolutionary inference methodology that focus on data integration in pathogen phylodynamics. These approaches include connecting sequence to trait evolution, but also the incorporation of covariates of the evolutionary process in the reconstruction procedures. I will discuss how these approaches allow testing hypotheses about the spatiotemporal spread of viruses. Specifically, I will focus on human seasonal influenza and the Ebola virus outbreak in West Africa.

The profile regression for spatial ecological studies

Aurore Lavigne (LEM-UMR CNRS 9221, Université de Lille, email: Aurore.Lavigne@univ-lille3.fr)

Summary: A major issue in epidemiology studies is to assess the individual effect of one factor on the incidence of a disease, when subjects are exposed to a mixture of correlated factors. In this context profile regression is becoming an attractive alternative to standard generalized linear models. Within the profile regression, individuals are clustered according their profiles of covariates thanks to a Dirichlet process mixture model, and a risk is attached to each cluster. Here we propose to adapt the profile regression to spatial ecological studies. In this context, to account for spatial dependence of residuals, a spatial random term is introduced in the regression. It is modelled with an intrinsic conditional autoregressive model. This model has been implemented within the Bayesian paradigm, and the R package PReMiuM has been modified to support this new spatial specification. Simulations show that the addition of the spatial term substantially reduces the mean square error of risk estimates. However, application to real datasets highlight some concerns as soon as the covariates are spatially dependent. Indeed the uncertainty about clustering increase. This is a joint work with Silvia Liverani, Marta Blangiardo and Anna Hansell.

Drug sales data analysis for outbreak detection of infectious diseases

Avner Bar Hen (MAP5, Université Paris Descartes, e-mail: Avner.Bar-Hen@mi.parisdescartes.fr)

Summary : Quantitative approaches are widely used to make predictions about the likelihood of an infectious disease outbreak, how the disease will spread, and how to control it. They are mainly based on clinical data that are collected through networks. Alternative source of data such as Internet data have been proposed among the last years. Drug sales data analysis is a tool that has been developed in the early 2000s for infectious diseases surveillance and outbreak

detection. These data present indeed many advantages for disease surveillance: the large volume and the completeness of the data, the rapidity to obtain the data automatically collected from pharmacies. The objective of this work was to determine the value of drug sales analysis for infectious disease surveillance in France. This is a joint work with Mathilde Pivette and Judith Mueller.