

## **Kolloquium „Statistische Methoden in der empirischen Forschung“**

Wann: 09. Dezember 2014, 17:00 – 18:30 Uhr

Wo: Lebenswissenschaftliche Fakultät, Humboldt-Universität zu Berlin,  
Hörsaal 2, 2. Etage, Invalidenstr. 42, 10115 Berlin

### **Brecht Devleeschauwer (Universiteit Gent)**

#### **Bayesian Estimation of True Prevalence from Apparent Prevalence in R: Introducing the "prevalence" package**

Diagnostic tests are crucial tools in many epidemiological studies. However, due to imperfect sensitivity and/or specificity, these tests provide biased estimates of the true prevalence. Based on prior knowledge about the test characteristics, however, it becomes possible to estimate the true prevalence from the apparent prevalence. To this end, a Bayesian approach is considered most appropriate, as it is able to flexibly incorporate such prior knowledge in the estimation process.

Different Bayesian models have been proposed in the literature to estimate true prevalence from apparent prevalence. However, the application of these models has remained cumbersome, as there are no tools available that implement them. Researchers have therefore been compelled to derive and develop their own models, most often in spreadsheet documents or in Bayesian software such as WinBUGS. This approach is error-prone, and hampers transparent and reproducible research.

To address this gap, we developed the 'prevalence' package in R, a free and open-source language and environment for statistical programming. As R, the 'prevalence' package is freely available and open-source. The package currently provides functions for the Bayesian estimation of true prevalence from individual samples tested with a single test; pooled samples tested with a single test; and individual samples tested with multiple tests. A variety of numerical and graphical diagnostics are available to assess model fit and convergence. The package further provides functionalities to establish prior Beta and Beta-PERT distributions for test sensitivity and specificity based on expert opinion. Work is in progress to develop functions for dealing with clustered samples and for including covariates in the different models.

By combining a variety of established methodologies for model fitting and evaluation, the 'prevalence' package provides a harmonized and comprehensive environment for true prevalence estimation. We therefore hope that the 'prevalence' package may become a useful tool for veterinary epidemiologists.