OMICS Group Conferences
Accelerating Scientific Discovery

2nd International Conference and Exhibition on
Biometrics & Biostatistics

June 10-12, 2013 Hilton Chicago/Northbrook, USA

Evidence for overdispersion in the distribution of Malaria parasites and Leukocytes in thick blood smears

Imen Hammami
Paris Descartes University, France

Background: Microscopic examination of stained thick blood smears (TBS) is the gold standard for routine malaria diagnosis. Parasites and leukocytes are counted in a predetermined number of high power fields (HPFs). Data on parasite and leukocyte counts per HPF are of broad scientific value. However, in published studies, most of the information on parasite density (PD) is presented as summary statistics (e.g. PD per microliter, prevalence, absolute/assumed white blood cell counts), but original data sets are not readily available. Besides, the number of parasites and the number of leukocytes per HPF are assumed to be Poisson-distributed. However, count data rarely fit the restrictive assumptions of the Poisson distribution. The violation of much of such assumptions commonly results in overdispersion.

Methods: We publish the first dataset on parasite and leukocyte counts per HPF. The data comprises the records of three TBSs of 12-month-old children from a field study of Plasmodium falciparum malaria in Tori Bossito, Benin. All HPFs were examined systemically by visually scanning the film horizontally from edge to edge. The numbers of parasites and leukocytes per HPF were recorded. The Pearson's test is used to check for overdispersion. Two sources of overdispersion in data are investigated: latent heterogeneity and spatial dependence. We account for unobserved heterogeneity in data by considering more flexible models that allow for overdispersion. Of particular interest are the Negative Binomial Model (NB) and mixture models. The dependent structure in data is modeled with Hidden Markov Models (HMMs).

Results: We found evidence that the Poisson assumptions are inconsistent with parasite and leukocyte distributions. Among simple parametric models, the NB model is the closest to the unknown distribution that generates the data. On the basis of model selection criteria AIC and BIC, the NB-HMMs provide a better fit to data than Poisson mixtures.

Conclusion: Undetected heterogeneity in parasite and leukocyte data may entail important misleading inferences when they are related to other explanatory variables (malariometric or environmental), so its detection is essential. The statistical dependence detected in parasite and leukocyte counts may bear critical information. An alternative PD estimation method that accounts for heterogeneity and spatial dependence should be seriously considered in epidemiological studies with field-collected parasite and leukocyte data.

imen.hammami@mi.parisdescartes.fr