UF UNIVERSITY of FLORIDA

"Survival Analysis and Phylogenetics in Infectious Disease Epidemiology" Symposium

Sponsored by

The Program in Research and Policy for Infectious Disease Dynamics (RAPIDD)
Fogarty Internationals Center, NIH

Emerging Pathogens Institute
Department of Biostatistics
Center for Statistics and Quantitative Infectious Diseases (CSQUID)

Wednesday, January 23, 2013
4 – 6 pm
2nd Floor Auditorium
J. W. Reitz Union, UF Campus
Reception to follow

Talks

First Presentation: Analysis of Viral Genetic Clustering

Victor De Gruttola
Professor and Chair
Department of Biostatistics
Harvard University

Second Presentation: Who infected whom? Molecular sequences reveal transmission patterns of emerging infections

Jacco Wallinga
Head, Unit Modelling Infectious Disease Epidemiology
Dutch National Institute of Public Health and the Environment

Victor De Gruttola

Analysis of Viral Genetic Clustering

Authors: Nicole Bohme Carnegie, Vladimir Novitsky, Rui Wang, Victor De Gruttola

Viral genetic linkage analysis will be essential for investigating HIV transmission dynamics and the effect of interventions on them; in particular it will be important to learn what host characteristics (e.g. disease status and demographics) are associated with genetically-linked infections. To illustrate the challenges in such analyses and to develop new methods to deal with them, blood samples from a household survey in the North East sector of Mochudi, Botswana were used to investigate whether rates at which subjects' viral genotypes cluster with others depends on their viral load levels (low/high) or ART status. Such clustering reflects HIV transmission dynamics; a tendency subjects with high VL to cluster more frequently with others might imply more transmission of viral strains prevalent in this group. A complicating factor in this analysis is that the probability of obtaining a sequence from a sample varies with viral load; samples with low viral load are more difficult to amplify. Naïve estimates of the group-wise probabilities of clustering based on observed clustering rates that ignore the differential rates of missingness are biased. We propose a novel approach for bias reduction based on bootstrap sampling. Sequences were obtained for 117 of 179 (65%) subjects with high viral load (>50,000 copies/mL) (HVL), 180 of 332 (54%) subjects with low viral load but not on ART (LVL), and 126 of 280 (45%)subjects who were on ART (ART). Two sequences were judged to cluster for this analysis if they differed in <12% of nucleotides. Preliminary results show that the probability of clustering between two individuals is highest if neither is on treatment, moderately elevated if one is on ART, and very low if both are on ART. After adjustment for relative population sizes all groups are more likely to cluster with LVL and less likely to cluster with subjects on ART. When we adjust for the presence of missing data, the tendency of subjects not to cluster with those on ART persists but the magnitude of the difference is substantially reduced**.** We conclude that bias in inferences regarding viral genetic clustering that arise from differential ability to genotype samples by subgroup can be reduced by appropriate methods for accommodating missing data.

Jacco Wallinga

Who infected whom? Molecular sequences reveal transmission patterns of emerging infections

An important question in studying infectious diseases is how transmission of infection is affected by the environment and by control measures. We address this question by first asking "who infected whom?" given observed case reports and molecular pathogen sequence data. We apply this approach to a dataset on a large outbreak of avian influenza A/H7N7 on poultry farms in the Netherlands in 2003, and we want to elucidate possible mechanisms of inter-farm transmission of influenza as these are largely unknown. The detailed genetic and epidemiological data allow us to answer the question "who infected whom?" and to reconstruct the time and direction of the transmission events. We find compelling statistical evidence for a statistical correlation between the direction of transmission and the prevailing wind direction at the time of transmission. These correlations are highly suggestive of a wind-mediated mechanism for avian influenza A(H7N7) in this outbreak that accounts for at least 18% of the observed spread of influenza.