ATELIER « ANALYSE DES DURÉES DE VIE ET DES DONNÉES HISTORIQUES D'ÉVÉNEMENTS »
16–19 MAI 2011

WORKSHOP ON "ANALYSIS OF SURVIVAL AND EVENT HISTORY DATA"
MAY 16–19, 2011

Risk prediction with complex studies

Tianxi Cai *
tcai@hsph.harvard.edu

The rapid emergence of new biological and genetic markers holds great promise for improving the prediction of disease risk and prognosis over time. To evaluate their utility in a clinical setting, a crucial step is to measure their predictive accuracy with prospective studies. However, it is often undesirable and/or infeasible to obtain marker values for the entire study population. Under such settings, the nested case-control (NCC) design is often employed as a cost-effective strategy. Under NCC design, markers are only ascertained for those who developed events and for a fraction of controls sampled randomly from the risk sets. The design, by employing incidence density sampling, generates a complex data structure and therefore a challenge for assessing the predictiveness of new markers. Existing statistical methods for analyzing NCC studies focus primarily on association measures. In this talk, we will discuss a class of root-n consistent inverse probability weighted estimators for making inference about predictive accuracy measures with data from NCC studies. We will show simulation results that suggest the adequate performance of the proposed procedures in finite samples. The new procedures will be illustrated with data from the nurse's health study to develop comprehensive risk prediction models for rheumatoid arthritis and evaluate the incremental value of biomarkers and a genetic risk score.

*Department of Biostatistics, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA.