広島統計談話会
Hiroshima Statistics Study Group

第 238 回談話会を下記のように開催致しますので
御参集下さいますよう御案内申し上げます。
You are cordially invited to the 238th meeting as scheduled below.

日　時：　2007 年 10 月 26 日（金）15:00～
Date：　October 26, 2007 (Fri) 15:00 -

場　所：　放射線影響研究所 講堂
Place:　RERF Auditorium

演　者：　スーザン M. ガイヤー（放射線影響研究所統計部副主任研究員）
Speaker:　Susan M. Geyer, Ph.D., Associate Senior Scientist, Department of Statistics, RERF

演　題：　「ある候補遺伝子があるときに生存と多重SNPの関連性を評価するための一つの方法」
Title:　"An approach for assessing multi-SNP association with survival in a candidate gene setting"

要　約：
Summary:

Background: There is little in the statistical literature on the application of large p/small n approaches to studies associating SNPs with survival. Thus, we have developed a multi-phase approach to assess the influence of multiple genes on patient survival and develop a prognostic model as appropriate.

Methods: A four-phase approach is utilized to build a multi-SNP model from a selection of SNPs

Phase I: Reduce pool of potentially prognostic SNPs to those with univariate Cox p<0.15
Phase II: Reduce SNPs identified in Phase I to 1 SNP per gene
a) Recode variable for each SNP so 1 = hazardous genotype(s) and 0 = protective genotype(s)
b) Exclude SNP if low minor allele frequency or low cell counts/number of events
c) Select one SNP from each resulting gene/high LD group.
Phase III: Perform parallel multi-SNP candidate model building techniques
a) Create score variable (1 df) for each of all possible combinations of the SNPs selected in Phase II. Fit each score variable in a Cox model. Group the models by number of SNPs included in score variable (i.e. 1 SNP models, 2 SNP models,) and rank by likelihood (-2LogL).
b) Include all SNPs from Phase II in a stepwise selection Cox model bootstrap (1K iterations) analysis. Determine the percent of iterations the SNP was included in the final model and rank the SNPs by inclusion percentage.

Phase IV: Determine final model.

a) Determine how many SNPs (M) to include in final model:

i) Plot -2LogL for the best fitting model for each number of SNPs

ii) Look at magnitude and gaps in inclusion percentages from Phase IIIb bootstrap analysis

b) Choose final model from list of models that contain M number of SNPs

c) Presentation of final model via Kaplan-Meier curves, Cox models, and time-dependent ROC curves

d) Permutation analyses to assess the significance of the model over expected findings from chance

Results: 73 SNPs from 44 candidate immune genes in 278 follicular lymphoma patients were modeled for overall survival. There were 59 events (21%); median follow-up of patients still alive was 59 months. Phase I analysis identified 17 SNPs in 12 genes. Phases II and III indicated a model with 4 SNPs (*IL8, IL2, IL12B, IL1RN*). Deleterious genotypes from these SNPs were adversely associated with survival ($p<0.0001$). Incorporation of clinical and demographic factors increased the effect and yielded a prognostic model (Kaplan-Meier $p=3.68 \times 10^{-13}$, 36 month AUC = 0.75) that compares favorably with current clinical prognostic indices (IPI/FLIPI) in this disease. Permutation analyses indicate the final model outperforms the best model from 85% of chance datasets.

Conclusion: This multi-phase approach identified 4 SNPs that are highly associated with survival in follicular lymphoma from a list of 73 SNPs. Validation of the model in an independent patient population is underway. Similar encouraging results using this multi-SNP modeling approach have been achieved in patients with diffuse large B-cell lymphoma.