REPRODUCIBILITY OF THE GENOTYPE-PHENOTYPE ASSOCIATIONS IN CASE-CONTROL STUDIES: STUDY DESIGNS AND STATISTICAL CONSIDERATIONS

K. Kopciuk[1], C. Zheng[2], C. Cass[3,4], R. Greiner[3,4], J. Mackey[3,4], S. Damaraju[3,4]

[1]*Alberta Cancer Board, Calgary, Canada*
[2]*University of Calgary, Calgary, Canada*
[3]*University of Alberta, Edmonton, Canada*
[4]*Alberta Cancer Board, Edmonton, Canada*

Email: *karen.kopciuk@cancerboard.ab.ca*

Association studies using gene polymorphisms (single-nucleotide polymorphisms or SNPs) are now widely used for predicting the susceptibility of populations to cancer and other polygenic diseases. Classical statistical methods to identify SNPs predictive of an outcome in an association study do not always perform well because of the large number of predictors relative to the number of observations. In the current study, breast cancer SNP data generated in the PolyomX Program at the Cross Cancer Institute, Edmonton were used to identify SNPs that are predictive of breast cancer risk and of association with estrogen, progesterone and Her-2-*neu* receptor status using classical statistical methods such as odds ratios and logistic regression (for cases only), and newer methods such as standard and Monte Carlo logic regression. Two small sets of cases (n = 174 and n = 164) and population controls (n = 158) were recruited over a short time span, which afforded the opportunity to identify significant SNPs in the first case-control data set and then validate the findings in the second data set. This study reports on the consistency of results between the two data sets, as well as between different statistical methods. Consistency among the receptors was also evaluated to see if the same SNPs are associated with positive or negative status. A larger case-control study (n= 750) is currently in progress to further validate the conclusions from this study.