

**PRE-ANALYSIS PROCESSING OF MASS SPECTRA:  
THE KEY ELEMENT FOR PROTEIN MARKER DISCOVERY**

Yutaka Yasui

*School of Public Health, University of Alberta, CANADA*

Matrix Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI TOF MS) has become a popular screening tool for discovery of biomarkers that are expected to have important translational potential in both therapeutic and preventive medicine. While many of the early results of mass-spectrometry-based protein biomarker discovery, especially in the field of early detection of cancer, have not been reproducible, some promising results have begun to appear recently. The success depends on a number of critical steps including: 1) epidemiological aspects of study design; 2) innovation in laboratory methods (e.g., sample preparation for maximizing signal to noise ratios of peptides/proteins that are of potential interest); and 3) proper data analysis methods for extracting the signals reproducibly and discovering differential signals associated with the phenotype of interest. Biostatisticians have important roles to play in each of these steps. In earlier biomarker discovery studies, data analysts focused largely on statistical learning methods and the high dimensionality of the data, without sufficiently ensuring scientific rigor in the critical steps above. Lack of attention to the above steps is likely to have contributed to the inability to achieve reproducible results. In this presentation, I will discuss some recent developments of MALDI-TOF MS pre-analysis processing methods for extracting the signals reproducibly. We illustrate how critical this analytic step is in the MALDI-TOF-MS-based discovery of potential protein biomarkers. The illustration will use multiple real-life examples including a large population-based cancer biomarker discovery project.