

Speaker: **Amit Phansalkar, Senior Research Scientist, Cytel Inc.**

Title: **Statistician's View of the Omics World**

Date: **Monday, May 18**

Time: **11 a.m.**

Location: **San Ildefonso Room, Research Library Study Center**

Abstract:

*Omics*

The world of omics (including genomics, proteomics, etc) has evolved as a way to measure biomarker activities and apply systems biology approaches to disease biology and drug discovery. A biomarker is defined as a parameter that is an objective indicator of biological or pathological or pharmacological processes in our body. In this talk we will look at the classification of biomarkers and review the technologies and statistical issues involved in biomarker discovery within the world of omics. A roadmap of challenges that is faced by the biotech world will then be discussed from a data modeling point of view. Two specific examples will be discussed to illustrate the implementation of Bayesian models to address these challenges.

*Reclassifying Genetic variants with uncertain clinical significance*

Clinical laboratories performing gene sequencing discover previously unreported and/or uncharacterized variants. Often these are missense or intronic mutations in which the contribution to disease cannot be predicted. Although family concordance studies show whether a variant segregates with disease in the family, the strength of evidence varies depending on the number and degree of relatedness of family members available for testing. A Bayesian model is investigated which accounts for the pedigree, inheritance patterns, and penetrance to determine the likelihood of a variant being a causative or deleterious mutation. Hereditary hemorrhagic telangiectasia (HHT) is used as a model for an autosomal dominant disease.

*HBM to accurately identify the presence of organisms from PCR sequencing data*

Management of prosthetic joint infections can be challenging especially when synovial fluid and intraoperative cultures are negative. The added value of performing 16S rDNA PCR and sequencing directly from surgical samples to detect pathogens infecting prosthetic joints will be evaluated. A hierarchical Bayesian model is developed to quantify the probability of contamination with each PCR result. After comparison to reference sequences in GenBank, the PCR sequences were identified based on a match score of greater than 97%. False positive rate and false negative rate in the dataset are estimated using the posterior from the Bernoulli binomial model. Using the model the sensitivity and specificity of PCR sequencing is compared with the results from culture as a gold standard.

Biosketch for Amit Phansalkar

I am a Sr. Research Scientist at Cytel Inc. where I lead consulting projects to design and provide analytical support to proof of concept clinical programs (early stage adaptive clinical trials) to reduce risk of product prospects for commercialization in biopharmaceuticals. I am also responsible for new product strategies and lead the development of customized simulation software products translating collaborative research projects to commercial products and services.

Prior to Cytel, I worked at ARUP, a diagnostic laboratory, where I lead the informatics team. Under my leadership the team provided informatics support to implementation of data mining models in genomics and proteomics resulting in increased efficiency and faster commercialization of diagnostic products.

I started my career as a biostatistician at Prolexys Pharmaceuticals Inc. and provided statistical support towards pre-clinical drug discovery projects. The work was primarily focused towards target discovery and validation in oncology. As a biostatistician at Myriad Proteomics Inc., I developed probabilistic models and applied systems

biology approaches for mining over 1TB of raw data generated from high throughput proteomics technologies to research the “human protein interactome”, a proprietary network map of human protein-protein interactions. My research has been recognized by several well cited publications, in leading scientific journals including *Nature*, *PloS Genetics* and *Journal of molecular diagnostics*.

I have a BS in mechanical engineering from University of Pune, India and MS in Statistics from University of New Mexico, Albuquerque NM.