Affiliates Workshop on Pharmacogenomics

February 12-13 , 2001

Abstracts

Pharmacogenomics and Molecular Medicine in Application to Rheumatoid Arthritis

Carl Edwards, Amgen

Collaborative studies between Inflammation Research, Clinical Development and the Expression Profiling and Proteomics Groups within Amgen Research and Development are addressing the use of microarray/expression profiling/proteomics of synovial tissues obtained from Rheumatoid Arthritic patients. Collaborations with internationally recognized Clinicians and Basic Researchers have been initiated since these interactions allow us to obtain target synovial tissue to assess the mechanism of action of Amgen compounds that are in clinical development. These studies will focus on three main uses of Amgen technology as it pertains to the Inflammation/Rheumatology Franchise at Amgen:

1. The identification of surrogate markers that can be utilized to determine if the drug is efficacious or not (i.e. something "better" than CRP and/or ESR).
2. The identification of patient subpopulations that may respond to one drug versus another, and
3. The identification of novel gene targets that can be used for drug discovery.

Experimental Approaches: We are currently involved in two Rheumatoid Arthritis studies which are in collaboration with clinical development, development and extramural research. Importantly, these studies will provide mechanism of action data on the protective effects of IL-1Ra and/or PEG sTNF-RI on bone and cartilage in chronically afflicted RA patients. The primary objectives of these studies are to confirm and to characterize changes in synovial biopsies to changes in joint architecture in RA patients. These studies have been designed similar to the recent Arava mechanism of action studies in synovial biopsy samples (Kraan et al. Arthritis and Rheumatism 2000). These data will also be correlated to X-ray and bone densitometry changes. The secondary objectives of these studies are to obtain ACR assessments, cytokines, bone and cartilage markers and differences in synovial tissue gene regulation using expression profiling on tissue samples that will be obtained longitudinally across a 52-week period.

These studies have the potential to provide value to Amgen Research and Clinical Development by establishing the "Proof of Concept" rationale for future Pharmacogenomic and Pharmacogenetics technologies to accelerate drug discovery and target evaluation.

Expression Profiling/Microarray/Data Analysis at Amgen

Dan Fitzpatrick, Amgen

The presence of a pre-existing genomic platform at the company has given Amgen a unique entrance into expression profiling using microarray technologies. The development of our current expression profiling platform and examples of utility will be presented, as well as the rationale for the maturation of the platform from a solely service-driven entity into a collaborative interdepartmental model.

High Throughput Statistical Quality Control and Statistical Analysis for MicroArrays

Kay Tatsuoka, Steve Clark, Jason Ruan, Mary Braewner,
David Gruben, Robert Knowlton and Frances Stewart
GlaxoSmithKline

Microarrays have arrived as a means of high throughput screening of genes and potential targets. Our lab processes hundreds of experiments per year. There are genuine outliers in the resulting data for both high and low signals due to for example dust or dropouts.Also, variability is inherent in all array-based gene expression data. This natural variation is due to differences in how genes respond to the specific experimental conditions of the array. Automated detection of outliers and quantifying the variability in a production environment is a requirement in microarray analysis for accurate measurement of low signals and in building high quality databases for further data mining.

We describe our methodology for identifying dropouts, outliers, and calculating p-values and confidence limits for gene expression measurements. We describe the role of individual spot quality indices in outlier identification as well as the use of replication. We describe our normalization procedure as applied to in-house datasets. We describe experiments that ensure that our confidence intervals are giving accurate coverage. We have found that our error model ensures that we are able to detect 2-fold changes with 99.5% confidence.

Inference of data from multiple cell-lines or multiple time points typically relies on a single

clustering and does not incorporate uncertainty in microarray data. We propose methodology to analyse these data that incorporates our error models and provides confidence statements about the classification of genes of unknown function. The value of the methodology is illustrated via simulations and on real datasets, and applied to k-means and hierarchical clustering.

Candidate Gene Analysis

S. Stanley Young
GlaxoSmithKline

The human genome project, gene chip technology and the desire of clinicians will lead to massive data sets where the question is: Which genes are associated with disease phenotype, side effects, and drug efficacy? There is a need for statistical methods to address this question. The potential problems are formidable. Analysis is likely to be of an exploratory nature as there are so many potential questions. There will be multiple testing dilemmas. Phenotypes will be described as multivariate observations and can arise by divergent mechanisms so the statistical methods need to be able to handle mixture data. Interactions are expected to abound. Our idea is to borrow and modify data mining methods to address these problems. Multiple testing will be addressed. Genotype/environment interactions will be addressed. Real data and simulated data will be used to explicate methods. Analysis software will be demonstrated. If we are successful, we can offer guidance to clinicians and help them prescribe "the right medicine for the right patient."

Genomic Simulations and Clinical Trials

Alan Menius, Dmitri Zaykin, S. Stanley Young, Meg Ehm and Micheal Mosteller
GlaxoSmithKline

One major strategy for the use of genetics data in the pharmaceutical industry is to collect and utilize genetic marker data from patients in ongoing clinical trials. It is believed that combining genetic marker data and clinical data will yield additional insights into both efficacy and safety. Little is known on how to plan a pharmacogenetic clinical trial so that genetic effects, if real, can be detected using available analytical techniques. One way forward is to produce simulated clinical populations where interacting genetic and clinical effects are estimated. These simulated data need to take into account genetic marker allele frequencies, linkage disequilibrium and their impact on clinical outcomes. These simulated populations can be used to help determine inclusion/exclusion criteria for a particular study and estimate the impact of allele frequencies, number of markers per gene, and linkage disequilibrium on power and sample size estimates.