Non-Clinical Statistics Conference | 2008

Organized by *the Adolphe Quetelet Society* (Belgian Region, IBS) in collaboration with the *Non-Clinical Statistics working group* of the German Region (IBS) and the *Biostatistics Section* of the Belgian Statistical Society

Abstract Book

Conference Theme
Statistical Methods for Pharmaceutical Research and Early Development

Non-Clinical Statistics Conference
Leuven, Belgium
September 23-25, 2008

Organizing Committee

Co-chairman NCS2008

Luc Bijnens
Vice President Quetelet Society (IBS)
Janssen Pharmaceutica (JJPRD), Beerse, Belgium

Bruno Boulanger
Biostats Section (Belgian Statistical Society)
UCB Pharma S.A., Braine L'Alleud, Belgium

Committee Members

Filip De Ridder (JJPRD-Janssen Pharmaceutica, Beerse, Belgium),
Ludwig Hothorn (Leibniz Universitaet Hannover, Germany),
Bernadette Govaerts (UCL, Louvain-la-Neuve, Belgium),
Philippe Lambert (UCL and University of Liege, Belgium),
Geert Verbeke (K.U.Leuven, Belgium),
Richardus Vonk (Schering AG, Berlin, Germany),
Ziv Shkedy (University Hasselt, Belgium).

Committee Secretariat

Katrien Van Laer (Janssen Pharmaceutica, Beerse, Belgium)
Kirsten Verhaegen (K.U.Leuven, Belgium)

Accountant Quetelet Socie ty

Hilde Wijnen (University Hasselt, Belgium).

Welcome to NCS2008!

On behalf of the organizing committees of this conference on Non-clinical Statistics, we have the pleasure to welcome you to Leuven, Belgium. Thank you all for the response you gave to this invitation! We hope that this meeting will be for each of you an opportunity to meet and share your experiences around stimulating scientific themes.

Several sponsors have given their financial support to the event. We thank them warmly for this support which will contribute to the success of this event: the Fond National de la Recherche, Fond voor Wetenschappelijk onderzoek, Arlenda, International Drug Development Institute, J&J Pharmaceutical Research and Development, SAS-Institute, UCB.

We wish you a pleasant stay in the city of Leuven !

Luc Bijnens and Bruno Boulanger

Social Events

City Trip and Drink at the Town Hall

On Tuesday September 23 at 17h30, immediately after the end of the short course of Prof. Davidian, we are organising a guided walking tour within the city of Leuven. Five professional city guides will show us the most famous, historically important and adorable spots and buildings during this 90 minutes guided walk through the city. At the end of the tour Minister of State Louis Tobback, the mayor of the city of Leuven, will welcome us in person at the town house for a drink until 20h00. It is an honour for us that the mayor will meet us in person in the city hall. He is one of the most famous socialist politicians of Belgium because he participated in several former governments as representative of the people in the parliament of Belgium. He has also been secretary of state of Belgium. All participants of the NCS2008 conference are invited to attend the tour and the drink generously offered by the city of Leuven. Please make sure to bring easy walking shoes for this city trip because the ancient pavements of the streets of Leuven are a challenge, particularly to ladies wearing high-heeled shoes. Please note that the city tour is immediately after the short course. You will not be able to stop at the hotel between the course and the city trip. Therefore we advice you to bring your travelling luggage to your hotel room before coming to the conference site.

Conference Dinner

All participants are invited to attend the conference dinner at Salons Georges. The dinner will take place on Wednesday September 24 from 20h00 to 23h00. The dinner will consist of a cold buffet to start, a warm buffet as main course, and dessert and coffee at the end. There are no additional costs since the dinner is offered by the conference organizing committee to all participants. Red and white wines of the house are included in the dinner.

Table of Contents

Scientific Program

Tuesday, 23 September 2008				
13:45	14:00	Welcome – Pre-Conference Workshop	Chair : Geert Verbeke Katholieke Universiteit Leuven, Belgium	
14:00	17:30	An Introduction to Nonlinear Mixed-Effects Models	Marie Davidian, North Carolina State University, Raleigh NC, USA	p 12
17:30		City tour of Leuven		
19:00		Drink at the Town Hall		
Wednesday, 24 September 2008				
08:15	08:30	Welcome – Main Conference	Luc Bijnens, Janssen Pharma, Beerse, Belgium	
Session 1		Model Based Drug Development	Chair : Filip De Ridder, Janssen Pharma, Beerse, Belgium	
08:30	09:15	Some principles of modeling and simulation in preclinical research and development	Philippe Jacqmin, Exprimo NV, Mechelen, Belgium	p 15
09:15	10:15	Contributed papers		
09:15	09:35	Using desirability indices for decision making in drug development	Didier Renard, Novartis Pharma AG, Basel, Switzerland	p 16
09:35	09:55	Adaptive model-based dose selection methods	François Vandenhende, Clinbay, Genappe, Belgium	p 17
09:55	10:15	Clinical relevance of dissolution specifications	Tom Jacobs, Center for Statistics, Hasselt University, Diepenbeek, Belgium	p 18
10:15	10:45	Coffee Break & Poster Session		

Session 2		Methodology I	Chair : Bruno Boulanger, UCB Pharma SA, Braine-L'alleud, Belgium	
10:45	12:05	Contributed papers		
10:45	11:05	Statistics in high-content biology	Rebecca Walls, AstraZeneca, Loughborough, UK	p 20
11:05	11:25	Combination of independent component analysis and statistical modelling for the identification of metabonomic biomarkers in H-NMR spectroscopy	Réjane Rousseau, Université Catholique de Louvain, Belgium	p 21
11:25	11:45	Expected and credible design space for analytical methods: a new perspective based on modeling and prediction	Pierre Lebrun, Université de Liège, Belgium	p 22
11:45	12:05	Symmetric and Asymmetric Sigmoid Curves: a close look at their statistical, numerical and mathematical properties	Charles Tan, Merck, West Point - USA	p 23
12:05	13:30	Lunch & Poster Session		

Session 3		Translational Medicine & Biomarkers	Chair : Luc Bijnens, Janssen Pharma, Beerse, Belgium	
13:30	14:15	Unified approaches for surrogate marker evaluation from multiple randomized trials	Geert Molenberghs, Censtat, Hasselt University, Belgium	p 26
14:15	15:15	Contributed papers		
14:15	14:35	Fit for purpose limits and tolerance intervals: connecting the biomarker assay performance to the clinical trial	Astrid Jullion, UCB Pharma SA, Braine-L'alleud, Belgium	p 27
14:35	14:55	Investigating association between behavior, corticosterone, heart rate, and blood pressure in rats using surrogate marker evaluation methodology	Abel Tilahun, Center for Statistics, Hasselt University, Diepenbeek, Belgium	p 28
14:55	15:15	Genomic biomarkers for a binary response in early drug development microarray experiments	Suzy Van Sanden, Center for Statistics, Hasselt University, Diepenbeek, Belgium	p 29
15:15	15:45	Coffee Break & Poster Session		

Session 4		Methodology II	Chair : Philippe Lambert, Université de Liège, Belgium	
15:45	16:30	Estimation of nonlinear mixed effects models in pharmacokinetics with the SAEM algorithm implemented in MONOLIX	France Mentré, Université Paris-7, France	p 32
16:30	17:50	Contributed papers		
16:30	16:50	A framework for estimation of area under the concentration versus time curves (AUC's) in complete and incomplete sampling designs	Thomas Jaki, Lancaster University, Lancaster, UK	p 34
16:50	17:10	Assessing repeatability and reproducibility of dose-response experiments	Marc Weimer, German Cancer Research Centre, Heidelberg, Germany	p 35
17:10	17:30	Mixed modeling using the SAS® PROC MIXED procedure: A simulation based approach to assess sample sizes and deal with a daily biostatistician dilemma for preclinical trials	Louise Baschet, Institut de Recherches Servier, Suresnes, France; Catherine Hessler, Sanofi pasteur, Marcy l'Etoile, France	p 36
17:30	17:50	Modeling Spatial Learning in rats based on morris water maze experiments	Christel Faes, Center for Statistics, Hasselt University, Diepenbeek, Belgium	p 37
20:00		Conference Dinner at Salon Georges		

Thursday, 25 September 2008				
Session 5		Chemical Manufacturing & Control	Chair : Richardus Vonk Bayer Schering Pharma AG, Berlin, Germany	
08:30	09:15	Application of the tolerance interval concept in quality assurance assessment of pharmaceutical products	Yi Tsong, Office of Biometrics, Center for Drug Evaluation, FDA, USA	p 40
09:15	10:15	Contributed papers		
09:15	09:35	Reconsidering shelf life: An update from the PQRI stability shelf life working group	James Schwenke, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, U.S.A.	p 41
09:35	09:55	Evaluation tolerance interval estimates: to capture or not to capture	Michelle Quinlan, University of Nebraska - Lincoln, Lincoln, U.S.A.	p 42
09:55	10:15	Development of statistical tools to test the equivalence between two measurement methods	Bernard Francq, Université Catholique de Louvain, Belgium	p 43
10:15	10:40	Coffee Break & Poster Session		

Session 6		Toxicology	Chair : L.udwig Hothorn, Leibniz University Hannover	
10:40	10:55	Semi- and non-parametric approaches to concentration-response modeling	Christian Ritz, University of Copenhagen, Denmark	p 46
10:55	11:20	Evaluation of the in vitro mutagenicity assays	Daniel Gerhard, Leibniz University Hannover, Germany	p 47
11:20	11:45	Invitrostat: An open-source R-GUI for the statistical evaluation of in vitro assays in toxicology	Frank Schaarschmidt, Leibniz University Hannover, Germany	p 48
11:45	12:30	Contributed papers		
11:45	12:00	The algae growth inhibition test – robust initial values for parameter estimation	Anke Schulz, Bayer Schering Pharma AG – Berlin, Germany	p 49
12:00	12:15	Assessing the similarity of bioanalytical methods	Jason Liao, Merck, USA	p 50
12:15	12:30	Dose-response evaluation using a combined parametric/non-parametric approach	John-Philip Lawo, Baxter BioScience, Vienna, Austria	p 51
12:30	14:00	Lunch & Poster Session		

Session 7		Omics	Chair : Ziv Shkedy, Censtat, Hasselt University, Belgium	
14:00	14:45	Tentacular analysis of high throughput omics data	Dhammika Amaratunga, Johnson & Johnson Pharmaceutical Research & Development LLC, Raritan, NJ, USA	p 54
14:45	16:05	Contributed papers		
14:45	15:05	Estimation of power and analysis of qPCR data with normal mixed models	Auli Partanen, Sami Hokkanen Orion Corporation ORION PHARMA, Turku - Finland	p 55
15:05	15:25	Order restricted clustering for dose-response microarray experiments	Adetayo Kasim, Center for Statistics, Hasselt University, Diepenbeek, Belgium	p 56
15:25	15:45	FARMS: a probabilistic latent variable model for summarizing Affymetrix array data at probe level	Djork-Arné Clevert, Institute of Bioinformatics, Johannes Kepler Universität Linz, Austria	p 57
15:45	16:05	A flexible probe level approach to improving the quality and relevance of affymetrix microarray Data	Chris Harbron, AstraZeneca, Macclesfield, UK	p 58
16:05	16:20	Closing and Concluding Remark	Bruno Boulanger, UCB Pharma SA, Braine-L'alleud, Belgium	

Non-Clinical Statistics Conference | 2008

Organized by *the Adolphe Quetelet Society* (Belgian Region, IBS) in collaboration with the *Non-Clinical Statistics working group* of the German Region (IBS) and the *Biostatistics Section* of the Belgian Statistical Society

Abstract
Short Course

An Introduction to Nonlinear Mixed-Effects Models

Marie Davidian, PhD[1]

In applications in the biopharmaceutical biological sciences, the time course of a continuous response for an individual may be well-described by a function that is nonlinear in one or more parameters, here an "individual" may be a human or animal subject, a laboratory sample, or some other observational unit. These functions often arise from theoretical, mechanistic considerations; for example, compartmental models for pharmacokinetic behavior for a single individual lead to systems of differential equations whose solution yields nonlinear functions describing the relationship between concentration and time for a single individual. Similarly, mechanistic models for disease progression, such as those characterizing within-host HIV dynamics, give rise to such nonlinear representations. In these settings, the model parameters often reflect directly individual-specific phenomena of direct interest, such as drug clearance or viral replication and death rates, and one key objective is to achieve an understanding of how meaningful such parameters vary across the population of individuals and the extent to which this variation is systematically associated with individual characteristics. For instance, knowledge of the form of the association between drug clearance and covariates such as renal function, weight, and age assist pharmacokineticists in developing appropriate dosing regimens. The nonlinear mixed-effects model is a natural statistical framework in which to address such objectives. In this short course, we introduce the conception and formulation of the nonlinear mixed-effects model. We then review popular approaches to inference under the model and popular software implementations. Extensions of the basic model are also discussed.

[1] North Carolina State University, USA

Non-Clinical Statistics Conference | 2008

Organized by *the Adolphe Quetelet Society* (Belgian Region, IBS) in collaboration with the *Non-Clinical Statistics working group* of the German Region (IBS) and the *Biostatistics Section* of the Belgian Statistical Society

Abstracts
Oral presentations

Session 1

Model Based Drug Development

Some principles of modelling and simulation in preclinical research and development

Philippe Jacqmin[1]

Since the introduction of the receptor theory in 1933 by Clark, modelling (i.e. mathematical conceptualisation) has continually been applied in biochemistry, physiology, pharmacology and therapy to describe, analyse and explain experimental observations. With the help of advances in both methods (e.g. in statistics: non-linear mixed effects modelling or Bayesian algorithms) and technology (e.g. in informatics: high speed personal computer), modelling has become more common place in drug development, essentially because the predictive performances of the models (when used in simulation mode) can be used to optimize the design/success of future explanatory or confirmatory studies or/and help in making decisions.

Advanced modelling and simulation has been used mainly in clinical development but it becomes clear that pre-clinical research and development can also benefit from application of the same approach. However, the uniqueness of pre-clinical development has to be recognized and the methodology adapted accordingly.

The main classical models that can be used in pre-clinical research and development will be reviewed (e.g. physico-chemical, enzyme, receptor, operational, PK, PBPK, PD, PK-PD (e.g direct and indirect response, link), K-PD, PK-PD-Disease, ...). The presentation will then discuss and provide examples of some implementation principles in pre-clinical development such as descriptive versus mechanistic models, in-vitro in-vivo correlation, rich versus sparse data, allometric scaling, PK-PD versus K-PD models, study design optimization, safety ratio and Bayesian approaches.

[1] Exprimo NV, Mechelen, Belgium

Using desirability indices for decision making in drug development

Didier Renard [1]

The clinical utility index (CUI) has been proposed as an integrated measure of clinical benefit/risk (Korsan et al, 2005). Its usage has focused on characterizing optimal dose ranges and selecting compounds when decisions are based on multiple attributes (e.g., safety and efficacy outcomes, quality of life benefits, drugability properties, etc.).
By definition the CUI is a weighted sum and requires defining utility functions to represent expected clinical value of possible outcomes.
In this talk, we discuss the use of the desirability index, DI, (Harrington, 1965) as an alternative measure to the CUI. Desirability concepts have been developped for the optimization of complex industrial processes and the weighted geometric mean is a popular choice in this field. A desirable feature of the geometric mean, which naturally translates to drug development applications, is that if one of the process' attributes is unacceptable, the process as a whole becomes unacceptable.
We will illustrate the dose optimization and compound comparison problems based on a real clinical application where the objective is to balance efficacy and safety outcomes. A simulation-based evaluation is performed to derive the distribution of DI while accounting for two sources of uncertainty: variability in estimated dose-response relationships and random variation in desirability functions, which are inherently subjective, in order to achieve a more robust assessment.
These principles can also be framed within a Bayesian decision theory setting where DI is a gain function and optimal decisions are determined by maximizing the posterior expected gain.

References:

Harrington E.C. (1965). The desirability function. *Industrial Quality Control,* 21, 494-98.

Korsan B., Dykstra K. and Pullman W. (2005). Transparent trade-offs – A clinical utility index (CUI) openly evaluates a product's attributes and chance of success. *Pharmaceutical Executive*.

---

[1] Novartis Pharma AG, Basel, Switzerland

Adaptive model-based dose selection methods

Francois Vandenhende and Julien Vanwinsberghe [1]

In this talk we will show how Bayesian modeling methods may be used to improve decision making and reduce costs in dose-ranging trials.
Our Bayesian toolbox includes adaptive design, evidence synthesis, predictive model and decision analysis solutions, all being available in a single software package called Decimaker. We will present the methodology and illustrate its value in a set of case studies targeting safety, biomarker and disease model responses.

[1] ClinBAY, Genappe, Belgium

Clinical relevance of dissolution specifications

Tom Jacobs [1] Geert Molenberghs [1] Filip De Ridder [2] Roel Straetemans [2] Luc Bijnens [2]

In Vitro-In Vivo Correlation modeling (IVIVC) is commonly used in preclinical and clinical biopharmaceutical research. It establishes a valuable link between the in-vitro dissolution and the in-vivo release of a drug product. Based on this link, the controlled release pharmacokinetic profile can be predicted from a subject's immediate release plasma concentrations time profile and the in-vitro dissolution time profile using the IVIVC model. If a controlled release capsule dissolves differently, this change in in-vitro dissolution properties can be translated into the corresponding altered in-vivo pharmacokinetic profile once an IVIVC is established.
In-vitro dissolution testing is the quality control test for a new batch of such a controlled-release drug product. If an in-vitro dissolution time profile falls within the specification limits at pre-specified time points, the batch is considered as acceptable, i.e., the new batch is expected to give the same exposure to the drug substance. However, the dissolution specifications are usually fixed very early in drug development, typically before first in-vivo testing and without the input of any IVIVC model. This raises the question of the relevance of these dissolution specifications. We combined the dissolution specifications with the IVIVC model to explore the e_ect of these specifications on the bioequivalence acceptance of a drug product. However, the drug concentration-time profiles are related to the clinical efficacy and safety of the drug product. Therefore, the IVIVC model can be combined with the pharmacokinetic-pharmacodynamic (PK/PD) model to understand the clinical relevance of the restraint of the in-vitro dissolution specifications.

[1] Center for Statistics, Hasselt University – Diepenbeek - Belgium
[2] Johnson and Johnson Pharmaceutical Research and Development, a division of Janssen Pharmaceutica – Beerse - Belgium

Session 2

Methodology I

Statistics in high-content biology

Rebecca Walls[1] **Chris Harbron**[1]

An important strategy in pharmaceutical research is to identify, as early as possible, compounds that are likely to fail at later stages of the drug discovery process, as a consequence of poor efficacy and/or toxicity to humans. One emerging approach is to employ a new technology, "multiparametric high content cell-based assays", which attempt to use *in vitro* cell models to mimic the complexity of the *in vivo* disease. Advanced imaging techniques are used to generate extremely large and complex datasets describing the response of a population of cells to a drug in a series of features, such as how the cells change shape or size, with the aim of building predictive models or "fingerprints" from the multiparametric assay data for well characterised compounds that elicit known responses. These fingerprints will then be applied to future compounds in order to predict the biological mechanism of action of the drug and its toxicity.

Traditional multivariate approaches are difficult to apply directly to such problems as the information captured for each multivariate feature is not a static data point; instead we have measurements taken over a dose range, resulting in a dynamic response to the compound for each of the features, yielding datasets with a three-dimensional cube-like structure (compounds by doses by features). In this presentation we evaluate and compare the properties of a range of statistical approaches for generating robust and reliable predictive models for future application from these three dimensional data.

[1] AstraZeneca – Loughborough - UK

Combination of Independent Component Analysis and statistical modelling for the identification of metabonomic biomarkers in ^{1}H-NMR spectroscopy

Réjane Rousseau[1] Bernadette Govaerts[1] M. Verleysen[2]

In order to maintain life, living organisms product and transform small molecules called Metabolites. The Metabonomics is a scientific platform, studying the development of biological reactions caused by a contact with a physio-pathological stimulus, through the metabolites. The 1H-NMR spectroscopy is used to describe the composition of the metabolites in form of spectra. Biologists can then confirm the development of a biological reaction if specific spectral regions ("biomarkers") are altered in regards to spectra obtained in physiological situations. However, this process supposed a preliminary identification in an experimental database of the biomarkers or spectral regions, to examine because of their changes in case of the biological response. Traditionally, this identification is realised, with some limitations, by examination of the 2 first components from a Principal Component Analysis. This talk presents a new methodology in two parts providing two kind of knowledge on ^{1}H-NMR metabonomics biomarkers: the identification of biomarkers and the visualization of the effects on the biomarkers caused by external changes of interest. A first part employs Independent Component Analysis in order to reduce the dimension. We present a process to decompose by ICA the spectral data into statistically independent components. We also propose solutions for questions specific to ICA concerning the number of components to estimate and their order. Comparatively to the usual PCA analysis, we expose the utilities of independent components to overview the data. The second part consists on statistical modelling applied to the results of ICA. We consider a panel of various statistical models adapted to the nature of the considered biomarker question.

---

[1] Université catholique de Louvain, Institut de Statistique
[2] Université catholique de Louvain, Machine Learning Group, DICE

Expected and credible Design Space for analytical methods: a new perspective based on modeling and prediction

Pierre Lebrun[1]

The Design Space (DS) of an analytical method is defined as the set of factor settings that provides satisfactory results, with respect to pre-defined constraints. The proposed methodology aims at identifying during optimization and validation phases a region in the space of factors that will likely provide satisfactory results during the future use of the analytical method in routine, through an optimization process of this method.

First, the DS is statistically defined as derived from the beta-Expectation Tolerance of prediction interval. Second, multi-criteria perspective is added in this definition as it is often required for optimizing most processes such as analytical method. Finally, a Monte-Carlo simulation is envisaged to numerically predict and identify the DS under uncertainty.

Examples based on high-performance liquid chromatography (HPLC) methods will be given, illustrating the applicability of the methodology.

---

[1] University of Liege – Liège - Belgium

Symmetric and Asymmetric Sigmoid Curves: a close look at their statistical, numerical and mathematical properties

Charles Y. Tan [1]

Four-parameter logistic curve has been widely used in quantitative immunoassays and bioassays, as well as concentration-response curves. However, alternatives are needed when data are inherently asymmetric. Richards model, a.k.a. "5PL", has been proposed, even promoted, by some in the scientific community as the alternative, despite the fact that it is long known in statistical literature that Richards model has poor statistical properties, e.g., large intrinsic curvature. We seek to demonstrate, via a real data set and bootstrap, that (modified) four-parameter Gompertz is a much better alternative: it is asymmetric, close-to-linear, and flexible enough to accommodate real data. The four-parameter Gompertz is also more stable numerically than the Richards model, which enables us to monitor and qualify assay runs through applying quality control methods on the parameter estimates. It is also important to point out that four-parameter Gompertz or Richards model, in fact, has a hidden 5th or 6th parameter, respectively, which is a binary parameter.

[1] Merck – West Point - USA

Session 3

Translational Medicine & Biomarkers

Unified approaches for surrogate mark er evaluation from multiple randomized trials

Geert Molenberghs [1]

The validation of surrogate endpoints has been initially studied by Prentice and Freedman. Noting operational difficulties, Buyse and Molenberghs proposed instead to use jointly the within-treatment partial association of true and surrogate responses, and the treatment effect on the surrogate relative to that on the true outcome. In a multi-center setting, these quantities can be generalized to individual-level and trial-level measures of surrogacy. Buyse and colleagues have proposed a meta-analytic framework to study surrogacy at the trial and individual-patient levels. Variations for various endpoints have been developed. Efforts have been made to converge to a common framework. This includes a so-called variance reduction factor and, importantly, an information-theoretic approach. Work has been done regarding sample size assessment, and the translation of the treatment effect on the surrogate to that on the true endpoint, leading to the so-called surrogate threshold effect.

[1] Censtat, Uhasselt, Belgium

References

Alonso, A. and Molenberghs, G. (2007). Surrogate marker evaluation from an information theorety perspective. Biometrics, 63, 180-186.

Burzykowski, T., Molenberghs, G., and Buyse, M. (2005). The Evaluation of Surrogate Endpoints. New York: Springer.

Fit-for-purpose limits and Tolerance intervals: connecting the biomarker assay performance to the clinical trial

Astrid Jullion[1] **Bruno Boulanger**[1]

The objective is to present the importance of the way to define the acceptance limits for an analytical assay and the way to assess the ability of the analytical assay to provide measurements within the acceptance limits. The impact of the acceptance limits (or fit-for-purpose limits) on the design, size and conclusions of clinical trials will be show in two cases with a first example in a bio-analytical assay that supports a PK study, such as a bioequivalence study. Secondly, the impact of fit-for-purpose limits will be shown on a Bayesian Dose-response adaptive design driven by a biomarker dosed using a bio-analytical assay.

We will propose the use of tolerance intervals to assess the quality of an analytical assay because it directly deals with the future accuracy of individual measurements, instead of with the past performance of the assay, as usually practiced.

The results will show that the acceptance limits of an assay and the way to assess its validity may have a major impact on the conduct of a trial and its conclusions. It also will be advised to integrate in clinical trials simulations the performance of the analytical method. Usually acceptance limits and validation rules definition are, so far, kept isolated within the analytical laboratories while they can have a major impact in the clinical trials they are supporting.

---

[1] Exploratory Statistics, UCB Pharma SA - Braine-l'Alleud - Belgium

Investigating Association Between Behavior, Corticosterone, Heart Rate, and Blood Pressure in Rats Using Surrogate Marker Evaluation Methodology

Abel Tilahun[1] John T. Maringwa[1] Helena Geys[2] Ariel Alonso[1] Leen Raeymaekers[2] Geert Molenberghs[1] Gerd Van Den Kieboom[2] Pim Drinkenburg[2] Luc Bijnens[2]

The drug development process involves identifying a compound and assessing its merit through rigorous pre-clinical and clinical trials. The pre-clinical stage is designed to assess the chemical properties of the new drug, as well as to determine the steps for synthesis and purification. In this stage of drug development, circumstances might dictate the use of alternative endpoints than the originally anticipated clinically relevant endpoint. In this regard, identification and evaluation of surrogate endpoints is of paramount importance. The validation methods enable to quantify degrees of association between the clinically relevant endpoint, also termed the true endpoint, and the alternative, surrogate endpoint. In this paper, we adapt the surrogate marker evaluation methodology of Alonso et al. (2003, 2006), developed for the case of two longitudinal outcomes, to the situation where either a longitudinal surrogate or cross sectional true endpoint or vice versa. The work is motivated by a preclinical experiment conducted to asses association between Corticosterone (CORT), heart rate, and blood pressure in rats. It was found that there is a weak relationship between CORT and behavior, and between CORT on the one hand and heart rate and blood pressure on the other hand, but a reasonably high degree of association was registered between heart rate and behavior.

Some Key Words: fractional polynomial; spline; surrogate endpoint, true endpoint; variance reduction factor.

[1] Center for Statistics, Hasselt University – Diepenbeek - Belgium
[2] Johnson and Johnson Pharmaceutical Research and Development, a division of Janssen Pharmaceutica – Beerse - Belgium

Genomic Biomarkers for a Binary Response in Early Drug Development Microarray Experiments

Suzy Van Sanden[1] Ziv Shkedy[1] Tomasz Burzykowski[1] Hinrich Göhlmann[2] Willem Talloen[2] Luc Bijnens[2]

Microarrays experiments have become a popular tool to examine the expression of thousands of genes at the same time. In early drug development experiments, the primary interest of researchers lies in finding genes that are differentially expressed when different condition (treatments) are tested. Recently, microarray data have also been considered as a means to select genes capable of serving as a biomarker for a primary response variable. Within this framework, one wants to assess the effect of a treatment on the response of interest by using information about the expression levels of a group of genes.

Shkedy et al. (2008) have considered a joint model for the gene expression and the response in pre-clinical experiments. The model allows to detect differentially expressed genes and to evaluate genes as biomarkers. In their study the response of primary interest was a continuous variable. In this paper we proposed a new model for biomarkers detection and evaluation for categorical response data. In particular, the response of primary interest is toxicity and the aim of the analysis is to find a subset of genes which can serve as biomarkers for toxicity. The method is applied to data containing information about toxicity levels of 38 subjects for whom microarray data for approximately 31000 genes are available.

Several approaches for gene selection are considered, amongst which an extension of the joint model to the categorical setting. We formulate a joint model for the binary response and the gene expression and we have shown that gene selection based on the joint modelling approach is similar to classical selection methods such as the BW ratio approach (Dudoit et al., 2002). The proposed method leads to the identification of a subgroup of genes (biomarkers) which are either treatment related, toxicity related or toxicity and treatment related from which toxicity can be predicted. In particular the new approach of identifying toxicity and treatment related genes allows investigators to focus on those genes that are influenced by the treatment and can discriminate between high or low toxicity levels. The new methodology and the software developed in this study will provide an important tool to investigators in the pharmaceutical industry for detection and prediction of toxicity using information from functional genomic experiments.

Keywords: Biomarkers; Microarray Experiments; Categorical Data; Joint Modelling.

---

[1] Center for Statistics, Hasselt University, Diepenbeek, Belgium
[2] Johnson & Johnson PRD, Beerse, Belgium

Session 4

Methodology II

Estimation of nonlinear mixed effects models in pharmacokinetics with the SAEM algorithm implemented in MONOLIX

France Mentré *[1] Marc Lavielle [2]

Population pharmacokinetics (PK) analyses are now used not only to provide mean estimates but also to make model selection, hypothesis testing, simulations and predictions based on all the estimated components: good estimation methods are therefore needed. The objectives of this talk are to describe and illustrate a new estimation methods based on a Stochastic Approximation EM (SAEM) algorithm as implemented in MONOLIX.

The statistical model for most population PK analyses is the nonlinear-mixed effects model (NLMEM). As opposed to linear mixed models, there are statistical issues to express the likelihood for these nonlinear models with respect to parameters so that first approximation methods (FO and FOCE) based on linearization of the model were proposed. It is well known that these methods have several methodological and theoretical drawbacks. They are also very sensitive to initial estimates which make lot's of run to failed to converge with a waste of time for the modeller.

EM algorithms are powerful algorithms for maximum likelihood estimation for mixed effect models, where random effects are defined as the missing data. However, again because the models are nonlinear, there is no simple extension for the E-step of the EM algorithm for NLMEM. Recently, stochastic extensions of the EM algorithms were proposed to perform maximum likelihood estimation (MLE) in NLMEM as the MCPEM algorithm [1] and the SAEM algorithm [2]. These algorithms avoid any linearization and are based on stochastic simulation for the E-step. The SAEM algorithm is based on well known statistical algorithm (MCMC, stochastic approximation, simulation annealing), its convergence to the MLE and its good statistical properties was proven for several extensions of standard NLMEM [3-5]. It is a powerful tool for MLE for very general incomplete data models

The SAEM algorithm is implemented in MATLAB in the free software MONOLIX [6] which handles several features that will be described. This software has been applied successfully to the analysis of several simulated and real data sets as for instance in [7], and examples will be shown. The different applications have demonstrated that stochastic EM algorithms are fast, rather insensitive to initial estimates and converge even for complex models where FOCE failed.

In conclusion, stochastic EM algorithms are new estimation methods needed in population PK analyses. Statistical research to extend the SAEM algorithm to handle complex type of data is ongoing and new releases of MONOLIX are scheduled incorporating the latest developments.

References:

[1] Bauer RJ, Guzy S. Monte Carlo Parametric Expectation Maximization method for analyzing population pharmacokinetic/pharmacodynamic data. In: D'Argenio DZ, editor. *Advanced Methods of PK and PD Systems Analysis*. Boston: Kluwer Acacemic Publisher, 2004: 135-163.

[2] Kuhn E., Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis*, 2005, 49: 1020-1038.

[3] Delyon B, Lavielle M and Moulines E. Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*, 1999, 27: 94-128.

[4] Samson A, Lavielle M, Mentré F. Extension of the SAEM algorithm to left-censored data in non-linear mixed-effects model: application to HIV dynamics model. *Computational Statistics and Data Analysis*, 2006, 51: 1562-1574.

[5] Donnet S, Samson A. Estimation of parameters in incomplete data models defined by dynamical systems, *Journal of Statistical Planning and Inference*, 2007, 137: 2815-2831.

[6] http://software.monolix.org/

[7] Lavielle M, Mentré F. Estimation of population pharmacokinetic parameters of saquinavir in HIV patients
with the MONOLIX software. *Journal of Pharmacokinetics and Pharmacodynamics*, 2007 34: 229-249.

---

[1] INSERM UMR738 and University Paris Diderot Paris, France

[2] INRIA, Orsay, France

A framework for estimation of area under the concentration versus time curves (AUC's) in complete and incomplete sampling designs

Thomas Jaki[1]

Nonclinical in vivo animal studies have to be completed before starting clinical studies of the pharmacokinetic behavior of a drug in humans. These pharmacokinetic studies are commonly analyzed using a two-stage approach. The first stage involves estimation of pharmacokinetic parameters for each animal separately and the second stage uses the individual parameter estimates for statistical inference. This two-stage approach is only applicable in complete data designs where each animal is sampled for analysis once per time point. In the case of rats and mice, where blood sampling is restricted, the batch design or the serial sampling design need to be considered. In batch designs samples are taken more than once from each animal, but not at all time points while in serial sampling designs only one sample is taken from each animal across all time points.

In this talk we will present a uniform approach to estimate the area under the concentration versus time curve (AUC) that is applicable to all three designs. Based on the asymptotic distribution we construct confidence intervals that are then evaluated against commonly used resampling based intervals. We will then show an extension to test for linear combinations of AUC's and illustrate its use in an example to test for deviations of dose proportionality.

[1] Department of Mathematics and Statistics, Lancaster University – Lancaster - UK

Assessing repeatability and reproducibility of dose-response experiments

Marc Weimerand Annette Kopp-Schneider[1]

In quantitative toxicology dose-response experiments are typically carried out to estimate a dose of a chemical compound that has a specific effect on a response variable. Statistical modelling of the data can be used to approximate the predictive relationship between dose and response and to identify doses of interest. Ideally, testing the same compound with a given assay produces nearly identical results even under conditions other than pure replication, e.g. different operators or laboratories. While in practice some variability is inevitable, demonstrating acceptably low variability is a critical step in the validation process of novel assays.

Statistical techniques for assessing repeatability and reproducibility of dose-response experiments can be linked to the statistical modelling approach, e.g. likelihood ratio testing of regression parameters, analysing random effects or calculating confidence limits for ratios of effective doses. A different approach is to call a given set of assay conditions an “observer”, the tested compound a “subject”, and the estimated effective dose an “observation”. Formulated in this way the large repertoire of descriptive tools and summary indices of the field of agreement statistics becomes available to assess repeatability and reproducibility of dose-response experiments.

In this paper we introduce this approach and show how agreement statistics can be used to assess endpoint variability of dose-response assays. Using simulations and real world data sets we compare this approach with regression model based techniques.

[1] German Cancer Research Centre – Heidelberg - Germany

Mixed models using the SAS® PROC MIXED procedure: A simulation based approach to assess sample sizes and deal with a daily biostatistician dilemma for preclinical trials…

Louise Baschet[1] **Catherine Hessler**[2]

In the non-clinical field, studies are often characterized by designs with longitudinal data, rather small sample sizes and potential missing data. Thus, the SAS® PROC MIXED procedure is commonly used to fit this data with linear mixed models. This procedure provides a huge variety of options, some of which are not well-known or are used without having a clear assessment of the impact on the results. In this context, we were interested in the influence of some of these options on type I error rate, power, and Information Criteria (AIC, BIC…) mainly for sample size calculation issue.

For this purpose, a standard two-factor design in parallel groups with repeated measures on time was selected. A limited number of time-points (3) was explored, but also a larger one (10) commonly encountered in non-clinical studies and rarely described in the literature.

A simulation approach was chosen to generate data corresponding to diverse experimental situations using various:

- types of interactions between group and time factors
- variance-covariance structures and correlation coefficients
- This data was analyzed through different modeling options:
- ML or REML methods
- different variance-covariance matrices from the one used to generate data
- the method for computing the degrees of freedom for the tests of fixed effects

General conclusions will be given to help the statistician in sample size calculation and modeling of data in a two-factor design with repeated measures, using the most appropriate options of SAS® PROC MIXED procedure.

[1] Institut de Recherches Servier, Suresnes, France

[2] Sanofi pasteur, Marcy l'Etoile, France

Modeling Spatial Learning in Rats Based on Morris Water Maze Experiments

Christel Faes[1] **Marc Aerts**[1] **Helena Geys**[2] **Luc De Schaepdrijver**[2]

In juvenile toxicity studies, it is of interest to test whether dosing juvenile rats with some compound of interest has an effect on its learning ability. The Morris water maze, developed by Morris (1984), is a behavioral experiment designed to test the spatial memory. In this experiment, a rat is placed into a circular pool divided in quadrants which contains a platform, hidden a few millimeters below the water surface. The rat is placed in the tank and must learn the location of the submerged platform through a series of trials. When repeating the experiment several times, the changes in time (latency) and distance (path) taken to reach the platform are indicators for the learning and memory abilities of the rat.

Analyzing data from the Morris water maze raises a number of statistical challenges. First, since the experiment is repeated at several time points, we are dealing with a longitudinal design. Second, when a rat does not reach the platform after 60 seconds, the rat was guided to the platform. This means that the outcomes of interest are right-censored. Further, multiple outcomes, of different nature, are of interest in this study (the latency is a time to event variable, while the path is measured as a count). The traditional analysis uses non-parametric tests to check for a possible dose-effect. However, due to the many tests performed, this approach lacks power. Also, these non-parametric tests do not account for the different complexities in the data. Here, an alternative method is proposed based on a generalized linear mixed model for outcomes of a different. This method accounts for the longitudinal design of the study, the right-censoring of observations when animals did not find the platform and the correlation between the time and distance taken to reach the platform.

Morris R (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neurosci Methods*, 11, 47{60.

[1] Center for Statistics, Hasselt University – Diepenbeek - Belgium

[2] Johnson & Johnson, PRD Biometrics and Clinical Informatics – Beerse - Belgium

Session 5

Chemical Manufacturing & Control

Application of the tolerance interval concept in quality assurance assessment of pharmaceutical products

Yi Tsong, Ph.D.[1]

Although the concept of tolerance interval and percentile estimation has been used in quality control for decades, many of the modifications of the approaches adopted in quality assurance assessment of drug products turned out to be both bias and obscured. In the recent years, FDA adapted the concept of tolerance interval and two one-sided tests in quality assurance of pharmaceutical products. This two one-sided tolerance intervals concept has been used in product specifications for uniformity of delivery dose, tablet unit dose content uniformity and in method transfer assessment. In the era of quality by design (QbD), products are sampled systematically with large amount; specifications are revised accordingly using tolerance interval concept. In this seminar, I will present the examples including unit dose content uniformity with small, medium and large sample sizes.

[1] Division of Biometrics VI, Office of Biometrics Office of Translational Science, Center for Drug Evaluation, FDA

Reconsidering Shelf Life: An Update from the PQRI Stability Shelf Life Working Group

James Schwenke[1] **Patrick Forenzo**[2]

The Product Quality Research Institute (PQRI) Stability Shelf Life Working Group was formed in late 2006 as a two-year effort. The Working Group members include statistical and pharmaceutical scientists from industry and academia. Our objective is to propose best practices for the analysis of stability studies and interpretation of quality attributes toward estimating shelf life consistent with the FDA Quality by Design (QbD) initiative. In doing so, our Work Plan provides for us to reconsider the definition of shelf life. Our current discussion are focused on three alternative definitions of shelf life:

(1) The shelf life for a degradant is the time that degradant stays within specifications.
(2) The shelf life for a degradant is the time when the true mean degradant level crosses the specification limit.
(3) The shelf life for a degradant is the time the degradant level for all tablets (or other unit) in the batch does not exceed the specification limit.

The Working Group's discussions are focused on the understanding that the shelf life is the parameter of interest. Shelf life should be a measure of overall response of a product, estimated on an acceptable proportion of samples that are within specification. This shifts the estimation of shelf life from the mean of the batch to a proportion of the samples within specification for each batch. The use of quantile regression, calibration techniques and interval estimation provide a direct estimate of shelf life with respect to storage time. Confidence, prediction, as well as tolerance interval estimates of shelf life are considered.

---

[1] Boehringer-Ingelheim – USA
[2] Novartis Pharmaceutics Corporation - USA

Evaluating Tolerance Interval Estimates: To Capture or Not to Capture

Michelle Quinlan[1] Walt Stroup[1] James Schwenke[2]

Statistical interval estimates are constructed to estimate parameters or quantify characteristics of a population. Various tolerance intervals have been proposed to capture a proportion of the observations corresponding to a distribution. Each interval has been theoretically developed. However, the estimates computed for each interval can differ greatly. To correctly interpret each interval estimate and make accurate inferences about the population, it must be determined how well each interval performs in capturing the desired proportion of observations. In addition, because tolerance intervals are being used with more frequency, a consistent definition of the intervals needs to be established.

A computer simulation was conducted to determine which interval estimates capture the middle proportion of the distribution of future observations, and whether the interval estimates are quantifying the same concept. Interval estimates considered include confidence intervals, prediction intervals, various tolerance intervals, simultaneous tolerance intervals, ß-expectation and ß-content tolerance intervals, and confidence/prediction intervals on confidence/prediction interval endpoints. Each interval estimate is evaluated and comparisons are made to determine the difference among the interval estimates. Clarification is made regarding whether capturing the middle proportion of the distribution is equivalent to capturing the middle proportion of the distribution of future observations. Simulation results are presented and interval estimates are quantified to establish the correct interpretation of each interval estimate based on what the interval actually captures, and determine the appropriate use of each interval across various situations.

---

[1] University of Nebraska -Lincoln - U.S.A.

[2] Boehringer Ingelheim Pharmaceuticals, Inc. – Ridgefield - U.S.A.

Development of statistical tools to test the equivalence between two measurement methods

Bernard Francq[1] **Bernadette Govaerts** **P. Pestiaux**

The needs of the pharmaceutical industry to quickly assess the quality of products and the performance of the manufacturing methods leads to the development and improvement of alternative analytical methods sometimes faster, easier to handle, less expensive or even more accurate than the method corresponding reference. These so-called alternative methods should ideally lead to results comparable to those obtained by a standard method known as a reference.

To test statistically the equivalence between two measurement methods, a certain characteristic of a sample can be measured by the two methods in the experimental domain of interest. Pairs of points (Xi, Yi), representing the measures taken by the reference method and the alternative one can be modelled by a linear regression (a straight line). The estimated parameters are very useful to test the equivalence. Indeed, an intercept significantly different from zero indicates a systematic analytical bias between the two methods of measurement and a slope significantly different from one indicates a proportional bias (the non equivalence between the two measurement methods).

The estimated parameters and their confidence intervals are then used to test the equivalence of the two measurement methods. To achieve this correctly, it is essential to take into account the errors in both axes. Different types of regression exist to handle these cases and a lot of confusion still exists in the literature. We review therefore the equations for estimating the regression by these different techniques (the ordinary least squares regression, the orthogonal regression, the least rectangles regression, the Deming regression, the bivariate least squares regression, the Bland and Altman Plot) and standardize the notations. Then, we'll focus on the most suitable equivalence techniques tests and we propose a new methodology to test the proportional bias and the analytical bias at the same time with a simultaneous confidence interval.

During the presentation, the different properties of these regressions, their advantages and disadvantages, will be analyzed and compared with simulations and real set of data.

[1] Université Catholique de Louvain (UCL) – Louvain-la- Neuve - Belgium

Session 6

Toxicology

Semi- and non-parametric approaches to concentration-response modelling

Christian Ritz[1]

This talk presents some of the results of a recent EU funded research project aiming at developing flexible concentration-response models for data from in vitro toxicology tests. The motivation behind the project is that the commonly used parametric concentration-response models for continuous and quantal data (eg the log-logistic model) may fail to describe the concentration-response data appropriately in some cases. In some application areas the proportion of problematic data may be as high as 20%. The inappropriateness of the commonly used parametric models may in some cases be remedied by using a more flexible class of semi-parametric concentration-response models, which are convex linear combinations of a parametric component and a non-parametric component. The parametric component is used to provide some basic "concentration-response" structure and it can be any parametric concentration-response model, whereas the flexibility of the non-parametric component is intended to capture all major departures from the parametric trend. This class of models is called model-robust semi-parametric models and includes as special cases in the extremes the fully parametric model and the purely non-parametric model. The models are fit in a two-step procedure, fitting each component separately. Subsequently a cross-validation procedure is used to determine the weight of the two components in the final model fit. Inverse regression methods are used for obtaining estimates of EC values of interest along with the corresponding confidence intervals. The performance of the proposed semi-parametric approach is illustrated in a case study and by means of simulations. The method is implemented as an extension package for the open source software environment R. A graphical user interface is also available.

---

[1] University of Copenhagen

Evaluation of the in vitro mutagenicity assays

Daniel Gerhard[1]

In many mutagenicity assays dose related trends of count or categorical data are of interest. Generalized linear models can be used to reflect the experimental design and account for covariates or other sources of extra variability between experimental units. For count data a quasipoisson or negative binomial distribution and for categorical data a quasibinomial or multinomial distribution are chosen, as these are common and overall quite robust assumptions.
For most applications it is hard to assume a certain dose-response curve; therefore a multiple contrast approach is proposed, which is not restricted to a specific trend alternative. By calculation of quantiles of a multivariate normal distribution with a correlation structure estimated from the data, approximate simultaneous confidence intervals are obtained for the ratio or odds-ratio of linear combinations of parameter estimates. The small sample performance is shown by simulation and the straightforward application of the proposed method is illustrated on real data examples.

[1] Leibniz University Hannover

An open-source R-GUI for the statistical evaluation of in vitro assays in toxicology

Frank Schaarschmidt[1]

Response variables in in-vitro toxicology comprise binary and count data as well as continuous variables which often exhibit heterogeneous variances or non-Gaussian distributions. The treatment and randomization structure of the experiments are often relatively simple factorial designs. Hypotheses of interest are typically two-sample comparisons, multiple comparisons to a control, comparisons to both a negative and a positive control, assessment of trends or proving equivalence to a control. Based on recent publications and a number of comparative simulation studies, a working group of Leibniz Universität Hannover provided a collection of statistical procedures to tackle these problems. Aiming to bring the available statistical methods close to the average user, we additionally provide an open-source graphical user interface (GUI) to the R-implementation of these methods. Based on Java, the GUI can both be installed locally and used in a Web application. It was developed in joint work with a group of the Fakultät Statistik, Technische Universität Dortmund. The GUI provides menus based on statistical as well as toxicological terminology and a documentation including a number of worked examples from different assays common in in-vitro toxicology.
Here, an overview of the implemented statistical methods is given and the application of the GUI is shown for a number of examples.

[1] Leibniz University Hannover

The algae growth inhibition test – robust initial values for parameter estimation

Anke Schulz[1]

The algae growth test, which investigates the growth inhibition of algae under addition of a substance, is a test routinely used in toxicology. The inhibition across doses can be described by a sigmoid model, but the widely-used probit model is unsuitable because of the limited range of values.
The four-parameter logistic model constitutes an adequate solution for the problem given the limited number of observations. A robust parameter estimation in form of the M-estimator proposed by Huber (Huber, 1964) can be used to cope with the large variation often encountered in the data.
Because of the non-linearity of the model, an iterative parameter estimation procedure has to be employed, which requires appropriate initial values to arrive at good estimates. Automated procedures for the identification of initial values found in the literature fail because of the large variations found in the data. Therefore, a new, robust procedure for the automated determination of initial values will be presented.

References:

- Huber, P.J. (1964). 'Robust estimation of local parameter'. *Annals of Mathematical Statistics* 35: 73-101 .
- Normolle, D.P. (1993). 'An Algorithm for Robust non-linear Analysis of Radioimmunoassays and other Bioassays'. *Statistics in Medicine* 12: 2025-2042
- Ritz, C. & Streibig, J.C. (2005). ‚Bioassay Analysis using R'. *Journal of Statistical Software* 12, Issue 5.
- Schulz, A (2006). 'Das vierparametrische logistische Modell und seine Anwendung bei einem ökotoxikologischen Problem'. Master Thesis. Humboldt-Universität zu Berlin, Germany.

[1] Bayer Schering Pharma AG – Berlin - Germany

Assessing the similarity of bioanalytical methods

Jason Liao [1]

A fundamental assumption of bioanalytical method (assay) is that the test sample behaves as a dilution or concentration of the standard sample. Thus, assessing similarity is a crucial part in a bioanalytical study. The common approach for assessing similarity is standard ANOVA analysis, which punishes the precise assay rather than poor assay. To overcome this problem, we propose a technique based on the idea of equivalence test but comparing the shape of curve directly. The new method measures the difference of response between the standard sample and test sample over concentration (dilution) levels and then justifies whether the difference is consistent by comparing them to the equivalence limit. The sensitivity and specificity of our method were investigated by the simulation study which shows that the new method does not suffer from the drawbacks of the conventional approaches.

[1] Merck

Dose-response evaluation using a combined parametric/non-parametric approach

John-Philip Lawo[1]

The Wessler-score (which is a seven-part ordinal scale) is often used as an outcome measurement to assess thrombogenicity. Thus, the thrombogenic potency of a drug can be evaluated by the dose-response relation between Wessler-scores and different dose levels.
Often the ED_{50} (i.e. the dose leading to an effect of at least 50% of the outcome) is used to characterize the dose-response curve. As the Wessler-score is an ordinal measurement, commonly used estimates based on the approximation by the normal distribution seem to be questionable for small sample sizes.

The poster will focus on estimating the ED_{50} by the E_{max} model when the responses are Wessler-scores. Three different types of confidence intervals (normal approximation, bootstrap-*t* and bootstrap percentile) are compared for different types of dose-response shapes and varying sample sizes.

[1] Baxter BioScience – Vienna - Austria

Session 7

Omics

Tentacular analysis of high throughput omics data

Dhammika Amaratunga[1]

(Joint work with Javier Cabrera, Nandini Raghavan, Jyotsna Kasturi, Willem Talloen, Luc Bijnens, Hinrich Göhlmann, and others)

One consequence of the tremendous advances being made in science and technology is that the scale at which multiple entities can be studied and/or measured and/or screened simultaneously has increased enormously. DNA microarrays, mass spectroscopy, high throughput screening, and molecular imaging are all examples of such technologies of relevance, and potentially of great importance, to the pharmaceutical industry. How to properly analyze and interpret the data these technologies generate remains a challenge for a number of reasons. Generally a combination approach is likely to be the most effective. Thus, for a standard comparative microarray experiment, a fairly typical analysis would include (1) an individual gene analysis to identify differentially expressed genes using a method that borrows strength across genes to increase efficiency (2) an analysis of gene combinations to identify affected biological processes and pathways (3) a meta-gene analysis to identify similarities/dissimilarities amongst the samples and the genes associated with any dissimilarities (4) an analysis of associated gene sequences to identify motifs of interest and perhaps (5) a network analysis to identify co-expression networks. We will review this multi-faceted approach along with an illustrative example.

[1] Johnson & Johnson Pharmaceutical Research & Development LLC, Raritan, NJ, USA

Estimation of Power and Analysis of qPCR Data with Normal Mixed Models

Auli Partanen[1] Eva Tas [1] Juha Akkila[1] Sami Hokkanen[1]

In recent years, a lot of attention is placed on microarray analysis. Thus importance of Quantitative Polymerase Chain Reaction (qPCR) is emphasized as a powerful, sensitive and reliable quantitative method for assaying gene expression in clinical and non-clinical studies. Statistical methods to analyze qPCR data are still improving since several factors need to be considered in the analysis: well-defined statistical model, proper internal control, amplification efficiency adjustment, data derivation and detection of outliers.

We will illustrate methods, results and power calculation in a qPCR study. Statistical analysis was conducted in parallel group design with repeated gene measurements using ??Ct method with and without efficiency adjustment. Power was estimated for expression ratio with selected type I error rate, standard deviation and number of measurements per treatment group.

Statistical analysis was performed using normal mixed model with fixed effects for treatment, gene and treatment by gene interaction, and random effects for sample and residual [1].

Calculation of power was based on the noncentral t-distribution and the log2 transformed expression ratio, which corresponds the ??Ct contrast of the four means [2]. Variance of the ?? Ct contrast was estimated simply as a sum of four equal variances.

References:

[1] Brown, H. and Prescott, R. (2006): Applied mixed models in medicine.

[2] Yuan, J., Reed, A., Chen, F. and Stewart, C. N. (2006): Statistical analysis of real-time PCR data. BMC Bioinformatics 2006, 7:85.

[3] Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. and Schabenberger O. (2006): SAS® for Mixed Models, Second Edition. Cary, NC: SAS Institute Inc.

[1] Orion Corporation ORION PHARMA – Turku - Finland

Order Restricted Clustering for Dose-Response Microarray Experiments

Kasim Adetayo[1] **Dan Lin**[1] **Ziv Shkedy**[1]

Dose-response microarray experiments consists of monitoring expression levels of thousands of genes with respect to increasing dose of the treatment under investigation. In the first instance, a gene-by-gene logistic model can be used to model the dose-response profiles.
However, this model will be unsuitable for genes with non sigmoidal profiles. Recently, Lin et al. (2007) discussed several testing procedures to test for monotone trend based on isotonic regression of the observed means (Barlow et al. 1972, Robertson et al. 1988). In the present paper, we propose an order restricted clustering method to find clusters of genes with similar dose-response profiles. Our proposal is motivated by the _-biclustering (Cheng et al. 2000). Order restricted clustering method offers potentials to uncover hidden patterns in a dose-response microarray data. This method can be used as a stand alone clustering method or as an exploratory tool for the existing methods.

References

[1] Barlow, R.E., Bartholomew, D.J., Bremner, M.J. and Brunk, H.D. (1972) Statistical inference under order restriction, New York: Wiley.
[2] Cheng, Y., and Church, G. M. (2000) Biclustering of expression data. In Proceedings of the 8th International Conference on Intelligent Systems for molecular Biology (ISMB'00), 93 - 103.
[3] Lin, D., Shkedy, Z., Yekuteli, D., Burzykowki, T., Golhman, H.W.H., De Bondt, A., Perera, T., Geerts, T., Bijnens, L. (2007a) Testing for trend in dose-response microarray experiments: comparison of several testing procedures, multiplicity, and resampling-based inference. Statistical Application in Genetics and molecular Biology, **6(1)**, article 26.
[4] Robertson, T., Wright, F.T. and Dykstra R.L (1988) Order restricted statistical inference, Wiley.

---

[1] Center for Statistics, Hasselt University, Diepenbeek, Belgium

FARMS: a probabilistic latent variable model for summarizing Affymetrix array data at probe level

Djork-Arné Clevert[1] Willem Talloen[2] Hinrich W.H. Göhlmann[2] Sepp Hochreiter[1]

High-density oligonucleotide microarrays, and in particular Affymetrix GeneChip arrays, are successfully applied in many areas of biomedical research. However, the large number of gene expression values, small sample sizes, and high noise levels lead to high false positive rates in extracting genes which are differentially expressed in different conditions. The false positive rate due to random correlations is a serious problem for biologists and medical researchers because if the significance level of their results is low or, even worse, they are misguided.
If the conditions are withhold in the first preprocessing step then random correlations between condition and expression values are considerably reduced in the second step. The first preprocessing step should exclude all genes, which do not contain a signal or are non-informative. As supervised feature selection approaches often suffer from overfitting, unsupervised feature filtering techniques are scarce and only look at the information content of the final expression value (signal intensity or non-Gaussianity). However for Affymetrix array data more information is available at probe level because a whole probe set records the expression value of a single gene. If probes of a probe set are governed by a common latent variable, then we associate this variable with the mRNA concentration and its variation with the mRNA variation, i.e. with the signal. Intuitively speaking, if probes of a probe set change synchronously across the arrays then this effect is very unlikely produced by noise and one should assume they are driven by a signal. Therefore we propose a probabilistic latent variable model to summarize array data at probe level and to assess the reliability of the probe set. Only such probe sets are filtered out where a variation of the latent variable can reliably be detected by a maximum a posterior optimization which combines and trades-off noise and signal likelihood.

Results:
We have evaluated FARMS on all public available spike-in data sets and at the Affycomp competition, where it outperformed all known summarization methods both with respect to sensitivity and specificity. Furthermore we found that our method excluded the non-informative probe sets without loss of sensitivity and specificity. The exclusion rates are about 99.5% while never losing a spiked-in gene in spiked-in data sets. As this objective technique results in non-informative feature reduction, it offers a critical solution to the curse of high-dimensionality in the analysis of microarray data.

---

[1] Institute of Bioinformatics, Johannes Kepler Universität Linz - Austria
[2] Johnson & Johnson Pharmaceutical Research & Development, a division of Janssen Pharmaceutica n.v. – Beerse - Belgium

A Flexible Probe Level Approach to Improving the Quality and Relevance of Affymetrix Microarray Data

Chris Harbron[1]

Microarrays have developed into a powerful tool in understanding biological processes at a fundamental molecular level through being able to simultaneously measure the expression levels of many thousands of genes. Unfortunately this great strength also gives rise to the greatest weakness of microarray technology, that of identifying false positives through multiple testing. This has led to many claims of discoveries which have failed to be validated in independent samples, leading critics to question the practical utility of microarray technology.

In a set of samples from a specific tissue, many genes will be unexpressed throughout and so will effectively just be contributing noise and false positives to any further analysis. An approach which would eliminate these unexpressed genes, whilst still retaining all the information containing genes would clearly be of great value. Various approaches have been applied, but they are limited by the use of arbitrary cut-offs and unrealistic simplifications about the behaviour of the technology and the underlying biology. In this presentation we will present an approach using Principal Component Analysis to identify those genes demonstrating consistent behaviour across the set of probes within an Affymetrix probeset, and demonstrate that restricting analyses to this set of consistent probes decreases the false discovery rate.

[1] AstraZeneca – Macclesfield - UK

Non-Clinical Statistics Conference | 2008

Organized by *the Adolphe Quetelet Society* (Belgian Region, IBS) in collaboration with the *Non-Clinical Statistics working group* of the German Region (IBS) and the *Biostatistics Section* of the Belgian Statistical Society

Posters

Using baseline gene expression for multi-Compound screening in early drug development experiments

Kasim Adetayo[1] **Dan Lin**[1] **Ziv Shkedy**[1]

The discovery of new active compounds is a big challenge for pharmaceutical research when differential response can be expected, like for oncology compounds. Here, we propose a method to analyse screening studies on oncology compound libraries on several tumor cell lines. The task is threefold. First, subsets of compounds that reveal some very distinct tumor growth inhibition profiles need to be identified . Second, cell lines need to be classified as either responding or not responding to the specific compounds. And third, baseline gene expression is used to differentiate responding from non-responding tumors in order to discover genetic signatures of response. This task is equivalent to simultaneous clustering of the compounds and cell lines in order to find groups of compounds that are active on responder cell lines and inactive on non-responder cell lines. We propose a method based on finite components mixture models to find clusters of compounds that are active on responder cell lines and inactive on non-responder cell lines. In addition, we suggest using supervised principal components and activity region finder to discover genetic signatures of response from baseline gene expression of the cell lines.

References

[1] Amaratunga D. and Cabrera J. (2004). Mining data to find subsets of high activity. Journal of Statistical Planning and Inference 122, No. 1-2, 23-41

[2] Bair E; Hastie T; Paul D; Tibshirani R. (2006). Prediction by Supervised Principal Components. Journal of American Statistical Society 101, No. 473, 119-137

[3] Gagarin A., Makarenkov V. and Zentilli P. (2006). Using clustering techniques to improve hit selection in high-throughput screening. Journal of Biomolecular Screening. 11, No. 8, 903 - 914

[4] Ma S. (2006). Empirical study of supervised gene screening. BMC Bioinformatics 7, 537

[1] Center for Statistics, Hasselt University, Diepenbeek, Belgium

Questions and Controversies related to Stability Modeling

Stan Altan[1] **Jyh-Ming Shoung**[1]

An overview of candidate models for the statistical characterization of stability in the context of a pharmaceutical drug product process is the starting point for raising questions about choice of model, poolability and inferences regarding expiration dating. In pursuing a choice of model, it is natural to consider batch as the independent statistical unit and specific questions related to error degrees of freedom, poolability across initials and strength are discussed. A dataset typically collected during development will form an example case study.

[1] Johnson & Johnson Pharmaceutical Research & Development LLC, Raritan, NJ, USA

Analysis of array CGH data for the detection of single-cell chromosomal imbalances

Michèle Ampe Geert Verbeke [1]

Based on Fluorescent In Situ Hybridization (FISH) analyses, previous studies on single human blastomeres have shown a high rate of chromosomal aneuploidy during early embryogenesis. Recently, it was shown that array CGH using BAC arrays enables the detection of chromosomal aneuploidies in single cells. To increase the resolution and the accuracy of the detection of chromosomal imbalances in single cells, we developed a finite mixture model enabling the detection of chromosomal and segmental aneuploidies.

The data are analyzed per single cell by fitting a finite mixture of Normal distributions per chromosome. Our finite mixture model consists of three Normal distributions corresponding to the duplication, the normal and the deletion group. The model corrects for the systematic bias, introduced by amplification, through an additional estimated clone-specific mean and clonespecific variance, both derived from a reference set of normal clones. The detection of aneuploidies is equivalent to searching for regions of successive clones that belong to one of the three groups. To obtain these regions, the posterior probabilities of the clones are smoothed by means of a loess smoothing. This approach has the advantages that loess smoothing is robust for outliers and the posterior probabilities take the a priori proportions of the three different groups into account.

The methodology was tested on single cells with a priori known abberations. Next, we analyzed single cells from three different cell types. The non a priori known aneuploidies that we detected, were confirmed by single cell SNP arrays. Using our methodology, we are able to screen single cells for genome wide copy number variations.

[1] Biostatistical Centre, Katholieke Universiteit Leuven, Belgium

Selection and Evaluation of Genomic Biomarkers for Depression

Dan Lin[1] **Ziv Shkedy**[1]

Microarray experiments have become an increasingly common laboratory tool which allows investigation of the activity of thousands of genes simultaneously and their response to a certain treatment.

Recently, microarray data have also been considered as a means to select genes capable of serving as a biomarker for a primary response variable. Within this framework, one wants to assess the effect of a treatment on the response of interest by using information about the expression levels of a group of genes.

In this study, we focus on a possible selection of genes to be biomarkers for clinical outcomes measured in a clinical trial. The outcome of primary interest is the depression level of patients before and after antidepressant treatment.

The aim of the analysis in this study is to identify and to evaluate genomic biomarkers which can be used for two purposes (1) prediction of the depression level of patients before treatment and (2) prediction of the response for treatment for the treated patients.

The modeling strategy for testing, validation and evaluation of genomic biomarkers is linked to the modelling approach for the evaluation of surrogate endpoints in randomized clinical trials and procedures for controlling the FDR are applied to address the multiplicity issue.

The method is applied to data containing information about depression levels of 200 patients for whom microarrays with 17,502 genes are available.

Keywords: microarray, biomarker, surrogate and true endpoints, FDR.

[1] Hasselt University, Center for Statistics, Diepenbeek, Belgium

Strategic roadmap for the IT support of the analysis and interpretation of data in drug discovery

Franky De Cooman [1] Roel Straetemans[2] Rudi Verbeeck[2] Luc Bijnens [2]

In many fields of pharmaceutical research statistics is only used to provide p-values. A fully data analytical approach, however, goes much further. Statistics can be integrated in the entire procedure to investigate scientific questions. When statisticians and other data analysts are involved in the planning of experiments, the data analytical part often becomes less complicated. In this paper we present an example of an automated statistical analysis resulting from this tight integration of discovery research, statistics and IT support. It gives the scientists the opportunity to perform their own data analysis because, after all, they know their data best.

This vision of integration of statistics and discovery research, facilitated by IT support, has strong impact on the researcher and the statistician. (1) We should make sure that the researcher's knowledge of statistics is sufficient to do basic analyses by offering a training path. These statistical courses highlight the pitfalls and explain the assumptions of the algorithms. We provide powerful yet intuitive tools like SAS/JMP and help-desk support so that scientists can analyse their own experimental data. (2) If statisticians can be freed up of routine analyses for the researcher, they could concentrate on designing new, more complex statistical solutions using advanced SAS tools (3) The statistical algorithms should be generalised and made robust, so that it can be applied to whatever datasets/variables on which the assumptions of the statistical method hold. It should also have an automatic model-building algorithm in order to provide the best model for the data.

The principle is illustrated in an analysis involving a generalized estimation equation (GEE) approach using GENMOD where CONTRAST and ESTIMATE statements need to be used. These statements are important tools when performing statistical analyses. For a statistician the creation of the correct code can however be cumbersome when dealing with many variables and/or different levels within variables. This paper will illustrate how the ARRAY statement can be used to circumvent these difficulties by letting SAS create the correct and necessary code.

[1] GBSC – Gentbrugge - Belgium

[2] Johnson & Johnson Pharmaceutical Research & Development – Beerse - Belgium

Classification of cell lines based on response to Treatment using PAMCHIPS

Arthur Gitome [1] Adetayo Kasim[1] Dan Lin[1] Suzy Van Sanden[1] Ziv Shkedy[1]

Peptide microarrays are based on methodology similar to that used in classical DNA microarrays, but with significant features like kinetic read-out. PamGene's three-dimensional PamChip microarrays offer a new alternative for drug discovery programmes. PamGene's technology platform can measure the effect of a drug on kinase activity in the patient's own tissue prior to treatment. For the present study, a phospho-tyrosine peptide profile using peptide arrays with a kinetic read-out is derived in lysates in the absence and presence of the kinase inhibitory compound. The objective is to find peptides signatures for the compound. Tumor cell lines used are known to be responsive or resistant to the oncology compound. To construct the classification rule, we considered combinations of feature selection and classification methods as proposed by Van Sanden, et al.(2008). The best result, as indicated by the lowest overall misclassification error, is obtained from the combination of tree based classification methods (bagging, boosting and random forest) and prediction strength.

References

[1] Van Sanden, S., Dan, L. and Burzykowski, T. (2008): Performance of Gene Selection
and Classification Methods in a Microarray Setting: A Simulation Study, Communications
in Statistics - Simulation and Computation, bf Vol. 37 , pages:418-433.
[2] Peptide microarrays in protein kinase and nuclear receptor research,
http://www.ngpharma.com/pastissue/article.asp?art=268321&issue=170

[1] Center for Statistics, Hasselt University, Diepenbeek, Belgium

Statistical evaluation of the local lymph node assay

Ludwig Hothorn[1]

In the Local Lymph Node Assay individual measured endpoints for each animal, such as cell proliferation, cell counts and/or lymph node weight should be evaluated.
The primary criterion for a positive response is when the estimated stimulation index is larger than an endpoint- and strain-specific relative threshold. When the lower confidence limit for ratio-to-control comparisons is larger than a relevance threshold, a biologically relevant increase can be concluded according to the proof of hazard. Alternatively, when the upper confidence limit for ratio-to-control comparisons is smaller than a tolerable margin, harmlessness can be concluded according to a proof of safety.
The related approaches are demonstrated by means of a real data example using the open-source software R.

[1] Leibniz University Hannover, Germany

A simulation based comparison of different statistical *k*-sample test procedures against ordered alternatives

Bernd Igl[1]

An elementary and challenging aspect in toxicological and drug development studies is to analyze the dose-response relationship of a chemical compound. To this end, classical parametric statistical mechanisms compare mean values to detect changes in response with increasing dose. In doing so, one assumes normally distributed data to obtain reasonable and valid results. However, in many cases data are not gaussian, but skewed, heteroscedastic, etc. such that common statistical test procedures may fail. In this talk, I compare different test approaches under different distributional scenarios via simulations.

[1] Bayer Schering Pharma AG, Berlin, Germany

Therapeutic Relevance of Bioequivalence Acceptance Limits

Tom Jacobs[1] Filip De Ridder[2] Sarah Rusch[2] Achiel Van Peer[2]
Geert Molenberghs[1] Luc Bijnens[2]

Pharmaceutical companies use pharmacokinetic measurements in bioequivalence (BE) trials as surrogate to prove that a new drug formulation or manufacturing procedure does not alter the safety and efficacy profile of the drug. In general, Health Authorities require that the 90% confidence intervals about the geometric mean test/reference ratios for both Cmax and AUC must fall between 80–125% to accept bioequivalence. For highly variable drugs and drug products, a high number of subjects is required to meet the current BE standards. Boddy [1] and Karalis [2,3] published approaches to correct or to widen, respectively, the acceptance ranges accounting for the degree of within-subject variability.

In 2006, Health Canada released a guidance on bioequivalence requirements for critical dose drugs [4] and proposed more stringent acceptance limits of 90-112% for AUC. In that guidance, critical dose drugs are defined as those where small differences in dose or concentration lead to serious therapeutic failures and/or serious adverse drug reactions.

In this work, the approach of Karalis was extended to adapt the BE acceptance ranges to the therapeutic window of the drug, quantified as the ratio of the Maximum Tolerated Dose/Therapeutic Dose and the Therapeutic Dose/Least Effective Dose. A series of simulations was carried out to assess the performance of the adapted acceptance range in a two-treatment, two-period cross-over study, with different sample sizes (12, 24 or 36 subjects), within-subject variabilities (15, 35 and 55%CV) and various ratios of MTD/Dose and Dose/LED. In addition, the method was retrospectively applied to theophylline data of Mistry [5], the digoxin data of Martin et al [6], and the phenytoin data of Meyer [7].

The results show that the approach has the desirable property of resulting in a more narrow acceptance range for doses near the boundaries of the therapeutic window and a wider acceptance range for products with a broad therapeutic window.

References:

Boddy A, Snikeris F, Kringle R, et.al. An approach for widening the bioequivalence limits in the case of highly variable drugs. Pharm Res 1995;12:1865-1868.
Karalis V, Symillides M, Macheras P. Novel scaled average bioequivalence limits based on GMR and variability considerations. Pharm Res 2004;21:1933– 1942.
Karalis V, Macheras P, Symillides M. Geometric mean ratio-dependent scaled bioequivalence limits with leveling-off properties. Eur J Pharm Sci 2005;26:54–61.
Health Canada Guidance for industry. Bioequivalence requirements: critical dose drugs. 2006

Mistry, G., Laurent, A., Sterrett, A., Deutsch, P.; ``Effect of indinavir on the single-dose pharmacokinetics of theophylline in healthy subjects."; J. Clin. Pharmacol.; 1999, Vol 39, 636--642.
Martin, D., Tompson, D., Boike, S., Tenero, D., Ilson, B., Citerone, D., Jorkasky, D.; ``Lack of effect of eprosartan on the single dose pharmacokinetics of orally administered digoxin in healthy male volunteers"; Br. J. Clin. Pharmacol.; 1997, vol 43, 661--664.
Meyer MC, Straughn AB, Mhatre RM, et al. Variability in the bioavailability of phenytoin capsules in males and females. Pharm Res 2001;18:394– 397.

[1] Center for Statistics, Hasselt University – Diepenbeek - Belgium
[2] Johnson and Johnson Pharmaceutical Research and Development, Janssen Pharmaceutica N.V. – Beerse - Belgium

An ordinal multinomial nonlinear model for the effects of investigational compounds on Taxol-induced mechanical neuropathy

Tom Jacobs [1] J. Adriaan Bouwknecht[2] Luc Bijnens[2] Geert Molenberghs[1,3] Roel Straetemans[2]

Taxol is an important drug in the treatment of cancer. It is a chemotherapeutic agent and is a cyto-toxic compound. It effectively damages malignant tissues but also affects peripheral neurons. Taxol induces a sensory deficit in rats after sub-plantar injection. The goal of the study is to find the counteractive properties of a number of compounds aga inst the peripheral neurotoxic effects of taxol.

Three compounds were tested in rats. These rats were treated once daily with a compound or vehicle subcutaneously for in total 8 days, whereas during the last 4 treatment days Taxol was administered in the left paw (sub-plantar). The neurotoxicity was assessed repeatedly over time by triple pricking the left paw up to 15 days after the first Taxol treatment. The reaction yields a multinomial ordinal response (0-3). The data were analysed using a cumulative logits model under the assumption of proportional odds (Agresti 2002). The resulting response profile indicated an increase of neurotoxicity for the first four days, followed by a recovery.

Rather than fitting a combination of two linear models during and after the 4-days of Taxol treatment, respectively, we used a sigmoid model (Gabrielson and Wiener, 2000). The plasma concentrations required in such a model are however unobserved. Therefore, a one-compartmental model was assumed. The regenerative effect of the compounds can as such be estimated from the E_{max} and EC_{50} parameters. A graph of the model predictions indicated a good fit of the data, however, the assumption of the unobserved plasma concentrations reduced efficiency of the parameter estimates.

References

Agresti, A. *"Categorical Data Analysis"*, John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2002.

Gabrielson, J., and Wiener, D. *"Pharmacokinetic and pharmacodynamic data analysis, concepts & applications"*, 3rd Edition, Swedish Pharmaceutical Press, 2000.

[1] Center for Statistics, Hasselt University – Diepenbeek - Belgium

[2] Johnson and Johnson Pharmaceutical Research and Development, Janssen Pharmaceutica N.V. – Beerse - Belgium

[3] Katholieke Universiteit Leuven, Biostatistical Center – Leuven - Belgium

Comparability of Three Statistical Methods to Assess the Reproducibility of Titer Based Assays during Validation

Eloi P. Kpamegan[1]

In the biopharmaceutical industry, it is not uncommon that standard deviation and %CV, which assumed that the distribution of the titers is normally distributed, are used to evaluate the reproducibility of titer based assays such as neutralization assays, serum bactericidal assay (SBA), Vero cell assays, etc. The problem stemmed from the absence of an appropriate numerical scale on which differing reproducibility can be quantified and objectively compared. Common practice in the biopharmaceutical industry considers the reproducibility of titers based assays to be acceptable as long as replicate titers remain within a two-fold range. Wood and Durham (1980) proposed as a measure of reproducibility "the probability that the maximum ratio of two distinct titers (obtained in the blind) on the same specimen will not exceed 2". The reproducibility titer based assays can also be effectively assessed using the percent of the results within one 2-fold dilution from the median titer (Median Method).

In this paper, reproducibility of titer based assays will be assessed and compared using three different methods: the Median Method, the %CV method, and the probability based method suggested by Wood & Durham (1980).

[1] Aeras Global TB vaccine Foundation – Rockville - USA

Gene Expression Analysis in Striatum after Acute Treatment with Antipsychotics

Setia Pramana[1] **Philippe Haldermans**[1] **Dan Lin**[1]

The study focuses on a dose-response experiment within a microarray setting, in which several microarrays are available for a sequence of increasing dose levels. Dose-response relationship was investigated in 6 different compounds.
The aim of the analysis is to identify genes for which a monotone trend of the expression levels with respect to dose can be detected. In particular, the parameter of primary interest is the increment of gene expression from baseline to the last dose level. We compare two different approaches: (1) a parametric approach based on non linear models and (2) isotonic regression. For each approach, marginal test statistics are calculated and used in order to rank the genes.
We discuss the similarity and the difference between the non linear models and the isotonic regression and in addition present an alternative approach in which ratio tests are used in order to rank the genes.

Keywords: Non linear models, Isotonic regression, tests for monotone trend, ratio tests.

References

Hu, J., Kapoor, M., Zhang, W., Hamilton, S.R., and Coombes, K.R. (2005) Analysis of dose response effects on gene expression data with comparison of two microarray platforms, *Bioinformatics, 21(17), 3524–3529*.

Lin, Dan., Shkedy, Ziv., Yekutieli, Dani., Burzykowski, Tomasz., Gohlmann, Hinrich., De Bondt, An., Perera, Tim., Geerts, Tamara., and Bijnens, Luc.,(2007) Testing for Trends in Dose-Response Microarray Experiments: A Comparison of Several Testing Procedures, Multiplicity and Resampling-Based Inference, *Statistical Applications in Genetics and Molecular Biology: Vol. 6 : Iss. 1, Article 26.*

[1] Center for Statistics, Hasselt University, Diepenbeek, Belgium

Direct Approaches to Shelf Life Estimation

Michelle Quinlan[1] Walt Stroup[1] James Schwenke[2]

Key Words: calibration, quantile regression, shelf life, out-of-specification, random effect

The goal of shelf life estimation is to determine the storage time when a product exceeds specification with some accepted probability. Current practices approach this indirectly through interval estimation about the mean change in response over storage time of stability limiting attributes. It is unclear if this approach captures the target probability. Alternatively, a more direct method of estimating shelf life as the parameter of interest is considered using quantile regression and calibration techniques. Quantile regression changes the focus of the regression model from the mean response to a specified proportion of the response population. Calibration techniques are used to predict the storage time associated with the intersection of the regression response with the specification limit. An interval estimate of shelf life is obtained based on the calibrated storage time.

Quantile regression and calibration methods are presented for estimating shelf life from multi-batch stability data where batches may be considered as fixed or random effects. Shelf life is estimated as the lower endpoint of a one-sided interval estimate of the storage time corresponding to the p^{th} percentile of the response distribution of a stability limiting attribute exceeding specification. This methodology allows shelf life to be directly estimated as the parameter of interest. While the current SAS® PROC QUANTREG procedure does not allow random effects, other ad hoc methods can be used to incorporate nonlinear quantile regression for random effects models. More appropriate quantile regression techniques incorporating random batch effects are being developed.

[1] University of Nebraska -Lincoln – Lincoln - U.S.A.
[2] Boehringer Ingelheim Pharmaceuticals, Inc. – Ridgefield - U.S.A.

Design and Analysis of Analytical Method Transfer Studies

James Schwenke[1] Dennis K. O'Connor[1]

An analytical method transfer study is a GLP study designed to successfully transfer technology of manufacturing, testing or assay processes from an origination laboratory to a destination laboratory. Following Quality by Design principles, a method transfer study is part of an analytical process to confirm the repeatability and ruggedness of the technology to be transferred. The primary objective of the study and the statistical analysis is to demonstrate equivalence between laboratory mean responses. Additional information demonstrating the consistency of analysts within laboratories and the proficiency of each laboratory and analyst to reproduce the expected result is obtained.

The typical method transfer study involves two laboratories with two analysts within each laboratory, testing samples over a two day period. The primary objective of the statistical analysis is to demonstrate that the laboratories produce equivalent results.

Secondary objectives are defined to demonstrate that the laboratories produce consistent results by demonstrating equivalence between analysts with each laboratory. In addition, each laboratory and analyst within each laboratory should demonstrate proficiency in producing acceptable results as defined by the target response. An equivalence testing strategy is employed to demonstrate the similarity of response obtained by the two laboratories. In the analysis of variance, laboratories and analysts within each laboratory are considered as fixed effects.

A testing procedure based on mixed model methodology is proposed to compare the variation in response between laboratories and analysts within laboratories. This proposed methodology indicates continued research being conducted to develop an equivalence testing strategy to compare one or more variances.

Depending on the pharmaceutical product involved, such as a capsule versus a solution versus a metered dose inhaler, different experimentation designs are necessary to accommodate the destructive sampling necessary for testing. Three designs will be discussed, representing independent sampling and variations on split and strip-plot designs. Examples of the analysis corresponding to each design will be presented.

[1] Boehringer-Ingelheim – USA

Analysis of Microaary Data in a Dose-Response Setting: Resampling Based Multiple Testing

Ziv Shkedy[1] **Dan Lin**[1]

The biotechnology of DNA microarrays allow the monitoring expression levels of thousands of genes simultaneously, and identifying those genes that are differentially expressed. As a result type I error (the probability for false identification) increase sharply when the number of tested genes gets large. In this talk we focus on a dose-response setting in which DNA microarrays are available for four dose levels (3 microarrays at each dose level). A gene is differentially expressed if there is a trend (with respect to dose) of the gene intensity. We discuss several approaches to test the null hypothesis of no dose effect versus an order alternative using isotonic regression. Resampling based False Discovery Rate (Benjamini and Hochberg, 1995, Ge et al. 2003, SAM et al. 2003) and resampling Family-Wise Error Rate (Westfall and Young, 1993) are used for controlling type I error.

Reference

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J.R.Statist. Soc*. B 57, 289-300.

Ge, Y., Dudoit, S. and Speed, P.T. (2003). Resampling based multiple testing for microarray data analysis. *University of Berkeley, technical report #633.*

Westfall, P.H. and Young, S.S (1993). *Resampling based multiple testing.* Willy.

[1] Limburgs Universitair Centrum, Center for Statistics, Biostatistics, Diepenbeek, Belgium

Visualisation of high dimensional data

Ziad Taib [1]

Visualising our data is often an important step towards a better understanding of the problem at hand. Sometimes, this step can even help visually and directly identify the underlying structure we are looking for.

However, when dealing with high dimensional data, it is often not obvious how to visualise the data in simple plots of lower dimensions. Gene expression data is but one type of data where we face this difficulty. There are many others.

Some of the most commonly used dimension reduction methods are often routinely applied in a naïve manner with poor or misleading results. We show some examples of such situations and propose alternatives based on recent technical developments in this area.

[1] Chalmers University of Technology – Göteborg - Sweden

Estimation of calibration interval for unknown concentration in a cDNA microarray experiment: Nonlinear mixed-effects models approach

Pushpike J Thilakarathne Geert Verbeke [1]

Currently calibration techniques are being used for the preprocessing of spiked in cDNA microarray experiments (Engelen et al., 2006). However, confidence intervals are of importance as to examine the uncertainty of the calibration estimate. We propose a calibration method for preprocessing spiked in microarray experiments in the context of nonlinear mixed-effects models.

Using the asymptotic properties of the calibration estimate we construct the 100(1 - α)% confidence intervals based on two different approaches. The first interval estimation method is based on the inversion of the regression function and the second involves inversion of asymptotic prediction intervals. A simulation study is performed to investigate the coverage probabilities of the constructed confidence intervals for a nominal level of 95%. It reveals that the interval estimation method based on the asymptotic prediction intervals is better. We applied this method to preprocess and to construct the confidence intervals for a dye-swap microarray experiment, based on data available in Hilson et al. (2004). We show that this method can be applied to more complex experimental designs like loop and interwoven.

KEY WORDS: Asymptotic; Nonlinear mixed-effects models; cDNA.

References

[1] Engelen, K., B. N. B. D. M. and M. Kathleen (2006). A calibration method for estimating absolute expression levels from microarray data. Bioinformatics 32, 490–5.

[2] Engelen, K. (2003). Normalizing microarray data. Bioinformatics 19(7), 893–4.

[3] Hilson, P. Allemeersch, J. and J. Beynon (2004). Versatile gene-specific sequence tags for arabidopsis functional genomics: 145 bibliography transcript profiling and reverse genetics applications. Genome Research 14, 2176–2189.

[1] Biostatistical Centre, Katholieke Universiteit Leuven, Belgium

Establishing Bioequivalence in Serial Sampling Designs

Martin Wolfsegger[1]

Pharmacokinetic studies are commonly analyzed using a two-stage approach. The first stage involves estimation of pharmacokinetic parameters for each analysis subject separately and the second stage uses the individual parameter estimates for statistical inference. This twostage approach is not applicable in sparse sampling situations where for each analysis subject only a single observation is available across all timepoints. This design is frequently used design in non-clinical in vivo animal studies.

This presentation will focus on establishing bioequivalence using ratios of area under the

concentration versus time curves (AUCs) in such serial sampling designs. Three new methods to test for bioequivalence via the confidence interval inclusion approach are developed and evaluated against commonly used resampling approaches.

---

[1] Baxter BioScience – Vienna - Austria

Author index