

# PSI Annual Conference

10-13 May 2015

**The Millennium Gloucester Hotel, Kensington, London**

## Programme of Abstracts





Promoting  
Statistical  
Insight

## PSI Annual Conference

10-13 May 2015

The Millennium Gloucester Hotel, Kensington, London

## Welcome Letter

# Welcome Letter

On behalf of the PSI Conference Organising Committee, I'd like to welcome you to London and The Millennium Gloucester Hotel for the 38th Annual PSI Conference. This year the conference theme is Relevant Applications in a Changing Environment so the agenda is bursting with hot topics and packed full with case studies. There are plenary sessions including estimands and sensitivity analyses, data transparency updates, joint HTA and regulatory advice and we are delighted to welcome Yannis Jemai, Vice President of Strategic Consulting and Product Management for Cytel, as our keynote speaker on Monday. The 3 days are a mixture of 5 plenary sessions, 18 parallel sessions and a total of more than 60 speakers to look forward to, not to mention a pub quiz and a barn dance for the energetic!

I therefore invite you to take advantage of all of the opportunities this conference brings with it; in meeting with old colleagues and friends, making new associates, learning something new, and above all, having fun! This is my first year as Conference Chair and I couldn't do it without the support of a fantastic committee so I would like to thank everyone who has

***“ The 2014 PSI conference attracted over 300 delegates and 20 exhibitors with particular high interest from senior positions within the UK and European companies ”***

been involved in the organisation of the conference. I would also like to take the opportunity to thank all of our exhibitors and sponsors as we would be unable to run this event without your continuing support. This year also sees the introduction of the Conference App so be sure to download it and keep up to date with the latest information.

After the conference we will be contacting you with a link to the electronic feedback form. Your feedback is very important to us in planning future conferences and we especially welcome ideas for the future or ways to further improve the conference to make it a better experience for you.

I look forward to meeting as many of you as I can this year, and wish you all an enjoyable and successful conference.

**Emma Jones, Veramed Limited**  
Conference Chair



## PSI Conference 2015: Collated Abstracts

### Monday:

#### Adaptive Designs: Past, Present, Future?

Our keynote speaker Yannis Jemai, Vice President of Strategic Consulting and Product Management for Cytel, will share his insight into the changing role of Adaptive Designs in our industry.

**Yannis Jemai (Cytel): Adaptive Designs: Past, Present, Future?**

#### Meta Analysis Panel Discussion

At last year's PSI conference, Sir Richard Peto stimulated a lot of discussion on the merits of random-effects vs. fixed-effects meta-analysis. Given the amount of heat generated then, we have invited a distinguished panel of experts in meta-analysis to continue the discussion and provide some answers on a range of topics. We have a selection of choice questions with which to turn up the heat again. In addition, we will be inviting questions from the audience. So if there was something you always wanted to ask an expert about meta-analysis or network meta-analysis, then here is a great opportunity to get answers.

Our panelists comprise Chrissie Fletcher (Amgen), Julian Higgins (Uni. Bristol), Armin Koch (Medizinische Hochschule Hannover) & Stephen Senn (CRP-Santé) - who will be ably and adeptly chaired by Byron Jones (Novartis).

These experts come from a range of backgrounds and will undoubtedly have strong (and differing!) opinions on the topics raised in this session. A lively debate and discussion is guaranteed - this is a session you won't want to miss.

#### Observational Studies: Big data: big deal?

Is big data really such a big deal? Statisticians typically have an innate and well-founded aversion to hype, especially with 'data science' fads. Our three speakers are probably no exception to this rule. However, despite the hype, 'big data' certainly has its place in pharmaceutical world, especially in the support of early- to late-phase clinical development. This session will focus on what 'big data' means to our speakers, where it is useful, and where it is not.

#### Speakers:

##### 1. Andrew Roddam (GSK): "How are we using big data in drug development – some perspectives from drug discovery through to development"

In this talk we will explore how pharmaceuticals companies are beginning to integrate "big data" into their drug discovery and development programmes. We will start our journey in the arena of discovery medicine and look at how the integration of Genetic with EHR data is starting to bring insights into novel pathways and endpoints. We will move forward through some of the more traditional uses of "big data" in Ph I/II and how this complements the clinical trial programme, before ending up discussing some of the more novel places where patient powered "big data" are starting to influence the way we think about and design our clinical programmes and some of the analytical challenges that result.

##### 2. Arlene Gallagher (MHRA): "Using electronic healthcare records to support clinical trials in the UK"

Trials are expensive. Anything that can reduce these costs is welcome and encourages more clinical trials to take place in the UK. Routinely collected electronic healthcare record (EHR) data could provide real efficiencies by optimising patient and site selection, removing the guesswork from clinical trial protocol planning and providing access to longitudinal data to supplement trial analysis. Arlene will describe how EHR data can be used to help in deciding where to site a trial and estimate the numbers of patients available for recruitment.

##### 3. Andrew Thomson (EMA): "Big Data – Challenges and Opportunities"

The analysis of increasingly larger data sets has changed the way that some drugs are developed, but not always in the way that might have been predicted. In this talk I will discuss how a specific application of this has changed the labels and requirements for registration for drugs defined by subgroups. I will also consider the statistical reasons why this application has been successful, and highlight on a key issue that statisticians, especially those involved in designing, analysing and interpreting clinical trials, are well-placed to address: the control of the Type I error rate.

The need for both statistical rigour, as well as the generation of new scientific insight may both be key to making the most from these bigger data sets

#### Pictures worth a thousand words: Innovative data visualisations

##### Speakers:

##### 1. Andreas Krause (Actelion): "A picture is worth a thousand tables: Visualization principles for clinical data"

A key aim of data and model visualization is the efficient display of the relevant information, enabling intuitive and accurate interpretation. The presentation establishes key principles of data and information visualization and provides illustrative case studies that implement the principles.

The principles of comparison as implemented in Trellis/lattice graphics are introduced. Graphical elements of graphics such as axes, lines, symbols, colors, legends, and three-dimensional displays are discussed and recommendations are given.

The presentation is based on the book chapter "Concepts and Principles of Clinical Data Graphics" in "A Picture is Worth a Thousand Tables" (Krause and O'Connell, Springer 2012).

##### 2. Chris Wells (Roche): "Visualization Techniques used to Display Growth during Tocilizumab Therapy for Polyarticular Juvenile Idiopathic Arthritis: 2-Year Data from a Phase 3 Clinical Trial"

The aim of the presentation is to show how to effectively summarise paediatric growth data and also show how to apply visualisation techniques to evaluate short- and long-term growth rates in paediatric patients. Data from the 2 year Cherish Study Results in patients age 2-17 with pJIA (paediatric rheumatic disease) and the potential impact on children's growth are presented. Growth data must be normalised prior to analysis due to the WHO age and sex dependent normal reference ranges. The presentation will use the Cherish data as an example to discuss:

- how to normalise, summarise and visualise the parameters of interest
- how to explore the baseline variables to identify potential factors affecting children's growth

### 3. Richard C. Zink (JMP Life Sciences, SAS Institute): “Subgroup analyses for personalized medicine”

Authors: Richard Zink, Russell D. Wolfinger - JMP Life Sciences, SAS Institute

In contrast to the “one-size-fits-all” approach of traditional drug development, the need to locate subjects with an enhanced treatment effect is a critical component for modern tailored therapeutics or personalized medicine. Typically, the goal is to identify patients receiving additional benefit from the treatment in terms of an efficacy response. Alternatively, finding subgroups based on important safety endpoints could be considered to determine those individuals experiencing a reduced risk of key adverse events, or to identify subjects for whom the new therapy may be inappropriate. Tree-based methods are naturally compelling in this context, and we review a few popular approaches that leverage recursive partitioning and hierarchical clustering. These analyses can be interpreted as finding the right patients for a given treatment. We compare them to optimal treatment regimes, which alternatively focus on finding the best treatment assignment (drug and/or dose) for each patient.

## Benefit Risk

To gain regulatory approval, a new medicine must demonstrate that its benefits outweigh any potential risks. Over the past several years, there has been a growing recognition amongst Industry Sponsor Companies and Regulators for the need of a more structured and consistent approach in assessing the benefit-risk balance of new therapies. This session will feature three talks exploring the potential for structured benefit-risk assessment to provide greater clarity of the benefit-risk balance to regulators, payers, and ultimately to patients.

### Speakers:

#### 1. Dr Shahrul Mt-Isa (Imperial College): “From qualitative to fully quantitative approaches to balancing benefits and risks of medicinal products for decision-making”

The evaluation of the balance between benefits and risks of drugs is fundamental to all stakeholders involved in the development, registration and use of drugs including patients, health care providers, regulators and pharmaceutical companies. Evidence on risks and benefits of drugs comes from diverse sources through the life-cycle of drugs. Clinical evidence is not the only important piece in benefit-risk evaluation; subjective judgements and preferences may also play a role. These pieces of information

are used to establish the benefit-risk balance of a medicinal product, whether qualitatively, partially-quantitatively or fully-quantitatively. The PROTECT Benefit-Risk Group clarifies this hierarchy of benefit-risk assessments with reference to the choice of methodologies and the complexity of the decision problems. In this talk, I will present the distinctions of the benefit-risk assessment hierarchy in decision-making, from qualitative to fully-quantitative approaches, through a case study example. By end of this talk, attendees will be able to identify where and when in the benefit-risk assessment process requires further quantifications and/or the use of more complex methodologies before a decision about benefit-risk balance can be made.

#### 2. Dr Alexander Schacht (Eli Lilly): “Structured Benefit-risk assessment: A review of key publications and initiatives on frameworks and methodologies by the EFPSI Benefit-Risk Special Interest Group (SIG)”

### Introduction

The benefit-risk assessment (BRA) of a pharmaceutical product interests various stakeholders throughout the life-cycle. The acceptance of standardized approach to BRA is rising and many examples are emerging. Statisticians need to play major roles in structured BRA within their organizations, and they can drive the shaping of future BRA, thereby having a deep impact on patients.

### Method

The EFPSI Benefit-Risk SIG searched for reviews and initiatives assessing BRA methodologies to assist those new to BRA in learning, understanding, and choosing methodologies. We summarize key points of the reviews and discuss their impact.

### Results

We provide introductory material, essential publications, and articles on special topics which were published between 2000 and 2013 to direct readers at various levels of expertise. Based on recommendations in these materials, we supply a toolkit of advocated BRA methodologies.

### Discussion

Although the acceptance of BRA is growing, the education on the benefits of BRA must continue to convince various stakeholders. This opens up opportunities, for statisticians in the pharmaceutical industry especially, to champion appropriate BRA methodology use throughout the pharmaceutical product lifecycle. Combining their methodological rigor and strong technical knowledge with

influencing skills, statisticians can lead benefit-risk assessments in order to contribute to sound decisions for the treatment of patients.

#### 3. Maria Costa (GSK): “Bayesian Benefit-Risk Assessment”

The Bayesian inference framework offers a tool for learning and updating one’s beliefs about particular parameters of interest. This aspect of Bayesian inference is especially attractive in the context of benefit-risk assessment, as existing information can be formally incorporated into the analysis of any emerging data. In addition, posterior probabilities offer a simple and clear device with which one may convey the benefit-risk balance to a non-statistical audience. This talk will present one approach which has been implemented internally to incorporate not only uncertainty in the observed data but also uncertainty at the parameter level through the use of prior distributions, and any potential correlations between benefit and risk endpoints

## Simulation of Trial Design

### Speakers:

#### 1. John Kirkpatrick (PPD): “Using Simulation to Improve Adaptive Trial Design”

Authors: John Kirkpatrick, Jürgen Hummel - PPD

We present two case studies showing how simulation can be used to inform decisions about adaptive clinical trial design. In the first example, simulation was used to decide the optimum number of interim analyses and their timing in a large Phase III non-inferiority study. In the second, we explored what was reasonable to expect from a small, first-in-man study using a variant of the Continual Reassessment Method. A good simulation plan is often iterative, and we discuss how the results of one stage of the simulation can be used to inform the design of the next. Issues associated with the implementation of methodology that is potentially unfamiliar to members of the project team should also be considered: problems may relate more to change management than to statistics.

#### 2. Alun Bedding (Roche): “Clinical trial simulations – an essential tool in drug development”

Authors: Alun Bedding, Nigel Brayshaw - Roche

The clinical development of investigational drugs is a complex and expensive process. The costs can be affected by decisions that are taken as clinical trials progress from one stage to the next (eg. dose selection studies and transitions from Phase 2 to Phase 3 clinical trials). Clinical trial simulations are being increasingly



viewed as an integral part of clinical development programmes and can be used to improve the understanding and decision making at every stage of drug development. The presentation will give the conclusions of a joint PSI/ABPI working group on simulations and resulting position paper.

### 3. Adam Crisp (GSK): “Simulating correlated cardiovascular endpoint data to assess power of different composite outcomes: an example using blinded data from an ongoing trial”

Cardiovascular outcomes trials have primary endpoints defined as composites of several individual components, such as time to first event of CV death, MI or stroke. The power to show a treatment benefit therefore depends on the extent to which there is a benefit on each of the component endpoints. We consider a scenario for an ongoing trial where it is hypothesised that an alternate composite endpoint might have more power than the originally defined primary outcome, with certain event types being common to both composites. A simulation is presented which explores the comparative power of different composite definitions, using a series of carefully partitioned blinded event sets as the foundation. By adopting the relative risk as a surrogate for the hazard ratio, a conditional Binomial framework is developed for the likelihood of individual events being observed in one treatment group vs the other, allowing for a presumed level of risk reduction for each event type. A re-sampling technique is then employed which generates composite outcomes that are drawn from the blinded data in such a way that accounts exactly for the underlying correlation structure, and enables power to be compared across a wide range of scenarios.

## Contributed Papers: Career Young Statisticians

This session is aimed at statisticians (presenters and attendees) with less than 5 years' experience working in the pharmaceutical industry.

### Speakers:

#### 1. Rhian Jacob (Roche): “A Career Young Statistician: A Rollercoaster 3 years involving Pharma, CRO and 90 mile commutes.”

My experience in the industry started with a placement year at Roche as a Biostatistician. I started out as a shy 20-year old with little work experience, average grades and terrible eye contact. I returned to final year with the confidence to speak out and

challenge others, and an eagerness to learn more.

After completing an MSc in Southampton I joined PPD in their Winchester office. I consciously wanted to develop SAS programming skills and gain broader experience by working on different therapeutic areas and clients. This was a great start to my career.

Almost two years later I was notified of a position available at Roche. By now I had established a strong network of friends in Hampshire and had become a homeowner with my partner. The Roche office is 90 miles away from my Hampshire home and the thought of accepting such a position was just crazy. But, crazy I am! I'm currently at Roche and taking advantage of the opportunities available to me, involving Bayesian designs, chairing an iDMC kick-off in Texas and attending the PSI conference.

This talk shares my experiences of both CRO and Pharma, and the challenges I've faced maintaining a work-life balance.

#### 2. Dan Lythgoe (Phastar): “A comparison of methods for survival modelling with a categorical latent covariate”

In clinical trials we are often interested in variables that cannot be measured directly; common examples include quality of life, depression and even tumour stage. We can use observed 'manifest' variables, such as questionnaire results, to make inferences about these latent variables. We may wish to measure the association between a latent variable and an outcome, for example survival time. However, incorporation of the manifest variables in a regression model is usually undesirable since a) each measures only one aspect of the latent characteristic, b) they can be highly correlated and c) there can be too many of them.

Latent class analysis is a multivariate method which uses discrete manifest variables to identify and characterise underlying categories of a latent variable. One method of estimating the effect of a latent variable on survival is to: 1) fit a latent class model, 2) assign patients to a latent class using the fitted model and 3) then incorporate the latent classes into a survival regression model as if they were observed. However, such multi-step approaches can result in biased parameter estimates (Bolck et al. 2004). Larsen (2004) presented an alternative, one-step method for survival models with latent class covariates.

We use data from a cancer trial to illustrate the differences obtained when tumour stage is treated as a latent variable and when it is treated as an observed variable. We also use simulated data

to compare the one-step approach with several variants of the multi-step approach.

### References:

- Larsen, K. (2004). Joint Analysis of Time-to-Event and Multiple Binary Indicators of Latent Classes. *Biometrics* (60) 85-92.
- Bolck, A., Croon M. and Hagenars, J. (2004). Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators. *Political Analysis* (12) 3-27.

#### 3. Ingrid Franklin (Veramed): “Using Established Genetic Risk Factors as Candidates for Melanoma”

Recent technological developments have allowed the facilitation of genome wide association studies (GWAS). A GWAS scans common genetic variants across the entire human genome with the objective of establishing whether any of these variants are associated with some particular disease or trait. A single nucleotide polymorphism (SNP) is a common form of genetic variation. Genome wide association studies produce a p-value that corresponds to each SNP and indicates whether it is associated with the disease or trait under examination. Due to large amounts of multiple testing and repetition during GWAS, a p-value can be required to be as small as  $5 \times 10^{-8}$  before a SNP is declared as genome-wide significant.

Intuition suggests that SNPs that are associated with some disease or trait are likely to also be associated with other 'genetically similar' diseases or traits. For example, one might expect a SNP that is associated with a pigmentation-related trait, such as fair hair, to also be associated with melanoma. This assumption can be examined by clustering p-values by disease area and comparing the cluster with randomly generated distributions using permutation testing. Exact randomisation tests are seldom performed in this research field due to the extreme size of data under analysis; the optimum number of permutations must therefore be carefully considered.

A case study will be presented which assesses similarities between SNPs from a GWAS on melanoma and SNPs that have already formally reached genome-wide significance. Differences between the significance levels of identical SNPs, will be examined under the hypothesis that significance is affected by geographical location and differing levels of UV radiation.

## Local HTA requirements

Title: Does the data speak for itself? Meeting the needs of HTA and reimbursement bodies in Europe

As the first stage to market access, a new medicinal product requires a marketing authorization from a drug regulatory agency based on three key factors: acceptable quality, safety and efficacy in a given patient population/clinical condition. The second stage, the so called fourth hurdle, is the assessment of the relative clinical- and cost-effectiveness of a new product and its value to the healthcare system within the context of current clinical practice. This assessment is to support pricing, reimbursement and coverage decisions and ensure that healthcare funding is spent appropriately. This assessment is generally performed by HTA agencies which give advice to national or local payers and other health care decision makers. Currently, the regulatory process for drug assessment is well established and the work of the International Committee on Harmonization, and use of the Common Technical Document have made the regulatory process for global drug development and assessment efficient. However, this is in contrast to HTA assessment process where there are fewer common processes and standards which can lead to marked differences in the specific evidence requirements and how the evidence will be viewed during assessments. Despite recent initiatives to identify synergies across these decision makers, there remains a challenge to product development teams to design and provide appropriate data to meet the differing requirements for relative clinical- and cost effectiveness assessment in different European countries. This includes not only generating the appropriate data within the clinical program but also presenting these data in a transparent and comprehensive way, often alongside other types of data, to provide an understanding of the applicability of the results to the local healthcare system. This session will identify some of the key issues which are critical to the perspective of these HTA decision makers but also explore how data can be generated to be fit for purpose and how the data are then viewed by the different HTA bodies.

### Chair

De Phung (Astellas)

### Speakers:

1. Jan McKendrick (PRMA Consulting Ltd)
2. Dr Karen Facey (Scottish Health Technologies Group)
3. Martin Scott (Numerus Ltd)

## Break-out Session: Subgroups

Following the success of last year's break-out sessions at conference, we are again running two round table discussion forums. The assembled audience at a session will be divided into groups, each group being given a focus and a list of suggested issues for their discussion. After a period of time for debating the key issues, the full audience reconvenes to hear the views of all groups. The theme for this first session is 'Subgroups'. Topics for discussion include: methodological approaches to interpretation of subgroups, regulation and the results of subgroup analyses, educating others on the limitations of subgroup analyses. This session is open to everyone, with or without prior subgroup experience. The focus is on interaction, idea sharing and discussion with peers and not lecture based.

## Treatment Switching

We are being asked more and more to follow our patients for longer and longer, by both health authorities and reimbursement agencies. This session aims to take a practical look, through three case studies of the analysis techniques that can be used to estimate treatment effects when patients are switching on to alternate therapies during the course of a trial.

### Speakers:

1. Elaine Wright (Roche): "The Trials and Tribulations of Treatment Switching: Practical experiences in Oncology of commonly used methods to adjust for treatment switching"

Authors: Elaine J. Wright and Iain Bennett

In randomized controlled trials, long term efficacy endpoints can be compromised by patients' crossing over or treatment switching (TS) before the event. TS can occur for many reasons including protocol defined switching after the primary surrogate endpoint is reached or when the active treatment is available as a subsequent line of treatment in clinical practice.

Methods developed in the 1990s to estimate the counterfactual treatment effect, confront the issue of bias found in some basic methods (e.g. censoring or excluding switchers). These more sophisticated methods have become increasingly useful in Health Technology Assessments where an estimate of the treatment's effectiveness over a life-time horizon is often required.

When TS exists in a clinical trial, the statistician's journey is broader than the application of the methods to the clinical trial data. The path the statistician takes goes from deciding if TS adjustment is required, appropriate, or even possible; to testing the assumptions and understanding the biases associated with the methods. This presentation will focus on the latter and share some ideas and practical experiences using these methods and testing the assumptions and biases underlying some of the commonly used methods.

### 2. Ioanna Gioni (Amgen): "Statistical Methods to Address Treatment Crossover in Randomised Clinical Trials"

Treatment crossover refers to the switching of participants in a clinical trial from their randomised treatment to another (either other arm or non-trial treatment) or to no treatment at all. The standard analysis of a randomised clinical trial (RCT) is the intention-to-treat (ITT) where participants are analysed according to their randomized treatment ignoring the treatment they actually received. Although the ITT analysis is the established method to evaluate the effectiveness of treatment policies it can provide biased estimates of the on-treatment effect in the presence of treatment crossover. Statistical methods such as inverse probability of censoring weights (IPCW), rank preserving structural failure time model (RPSFTM) and iterative parameter estimation (IPE) can deal to an extent with treatment crossover. This presentation aims to describe the basic principles of these methods. Results from the EVOLVE (Evaluation of Cinacalcet Therapy to Lower Cardiovascular Events) trial where these techniques were applied to explore the impact of treatment crossover on the on-treatment effect cinacalcet on a composite endpoint (consisting of all-cause mortality and major cardiovascular events) in patients receiving hemodialysis with moderate to severe secondary hyperparathyroidism (sHPT) will be also presented.



**3. Heinz Schmidli (Novartis): “Analysis of clinical trials with recurrent events endpoint and treatment switching”**

Recurrent events endpoints are important in many therapeutic areas, such as multiple sclerosis (relapses), asthma or COPD (exacerbations), gout (flares), and epilepsy (seizures). Clinical trial designs in these areas may involve treatment switching. For example many clinical trials consist of a core phase and an extension phase. In the core phase, patients are randomized to one or more regimens of the experimental treatment or to placebo. In the extension phase, patients are then switched to one of the experimental arms. A joint analysis of the core and extension data can provide valuable insights on the possibly time-varying effect of the experimental treatment. We discuss flexible statistical models to evaluate such data, and use a clinical trial in multiple sclerosis to illustrate the methodology.

**Reference:**

Chen Q, Zeng D, Ibrahim JG, Akacha M, Schmidli H (2013)  
Estimating time-varying effects for overdispersed recurrent events data with treatment switching. *Biometrika* 100(2):339-354.



## PSI Conference 2015 Abstracts for posters

### A Bayesian approach to the 3-parameter Emax model for the assessment of dose response and dose comparison

Toby Batten

CMed

The Emax model is now a well-established technique for assessing the dose response relationship for a new drug during early phase clinical trials. Through Emax modelling it is possible to estimate the maximal treatment effect, the dose which produces 50% of the maximal effect and the placebo effect.

Bayesian analysis enables us to incorporate historical data into our statistical models. The availability of historical data gives justification for a reduced sample size. In early phase clinical studies this is often not applicable to active treatment groups, however it can be applied to the control group, especially in disease areas where a number of clinical trials have already been performed.

The SAS MCMC procedure can be used to perform this type of modelling. Through MCMC not only can we calculate the terms of the Emax model but we can also adjust for any additional effects and compare the posterior samples to establish individual dose effects and differences. Therefore this approach can give comprehensive insight into drug efficacy as well as identifying the most effective dose.

### Pattern Mixture Modelling Approaches for Time-to-Event Data

James Bell, Simon Fink

Boehringer-Ingelheim

Non-informative censoring is required as an assumption for most common time-to-event analysis techniques used in the analysis of clinical trials, including Kaplan Meier analysis, Cox regression and the log-rank test. However, it is also a strong assumption that is not likely to be that realistic in many cases.

Here, we propose an implementation of pattern mixture modelling in time-to-event data as a framework for implementing a range of sensitivity analyses for informative censoring. In doing so, we build upon ideas presented at the 2014 PSI Conference by

O'Kelly. The method uses Kaplan Meier imputation as described by Taylor et al. (2002), with modifications to introduce clear assumptions regarding behaviour after censoring. In particular, we look at application of delta-adjustments to adjust for a worsened outlook after censoring and reference-based methods to account for treatment switching/discontinuation. Finally, we outline how patterning of the data set (e.g. by reason for censoring) may be combined with these techniques to perform more complex sensitivity analyses and address alternative clinically-relevant estimands. It is anticipated that a tool to implement these methods will be made freely available.

#### References:

O'Kelly M, Lipkovich I; 2014 PSI Conference presentation: "Using Multiple Imputation and Delta Adjustment to Implement Sensitivity Analyses for Time-to-Event Data".

Taylor J M G., Murray S, Hsu C; Statistics and Probability Letters 2002, 58 221-232: "Survival Estimation and Testing via Multiple Imputation".

### Eliciting expert opinion to improve decision making in clinical drug development

Nicky Best, Nigel Dallow, Tim Montague

GSK

For the past 12 months, GSK has been piloting the use of prior elicitation techniques to enable quantification of existing knowledge in the absence of directly relevant data, and to help predict probability of success of next study(s) at key milestone decision points for all phases of clinical drug development. This initiative forms a key component of an R&D-wide focus on innovation in clinical design at GSK, which aims to establish Bayesian approaches and use of prior distributions as standard practice to support internal decision-making and analysis.

In this presentation, I will give an overview of the prior elicitation process, and discuss some of the benefits and challenges we have experienced from using prior elicitation techniques at GSK. Key issues include the pros and cons of aggregating priors from several experts versus retaining the individual priors, and how to manage the tendency for over-optimism that is inherent in many experts' priors. I will discuss some recent work to address the

latter problem by eliciting a mixture prior: experts are first asked for their judgments that the drug will 'work', and are then asked to elicit a conditional prior for the treatment effect assuming that the drug works. The presentation will be illustrated using various case studies ranging from POC to Phase 3b studies.

### Predicting the date when the nth event will occur – Are Statisticians also Wizards?

Sandrine Cayez

CMed

Have you ever been asked to predict when x number of subjects will be randomized or when the Interim analysis (IA) will occur? Well if you have not yet you will be asked soon and who knows it may be tomorrow.

In our world, cost reduction and reduction of a trial duration (to be able to submit to regulatory agencies sooner) are key drivers. A better planning of study conduct is therefore required and rely on predicting when the recruitment will be completed or if applicable when an IA will occur.

In the past I have seen Project Managers predicting the recruitment and trying to determine when the nth event will occur (for an IA) using Excel spreadsheets and past experience alongside with the study clinical data, but should we as statisticians be not more proactive and use our vast statistical knowledge to help them (surely we can so better than a regression line in Excel)?

The statistical methods we can apply to this very important problem do not have to be complicated and require a large set of assumptions or lengthy computations as we all know that no clinical trial goes according to plan! Prediction must be adjusted several times during the study based on the current data or additional input (new sites selected and opened) so know your study.

Using simulation in SAS and known distributions for survival data types (death, Progression Free Survival, discontinuation, etc.) with maybe just a sprinkling of Bayesian can enable us statisticians to provide the team (and the stakeholders – never forget them) with the estimated date of interest and the most important part - confidence intervals!



## PhUSE Good Programming Practice Working Group: Providing industry standard to make it easier to share and validate programs

**Shafi Chowdhury**

**Shafi Consultancy**

Good Programming Practice (GPP) has long been a challenge for programmers within the Pharma industry. As the industry becomes more mobile, more international, the importance of this cannot be stressed enough. However, as the industry changes, so the role of programs from statisticians also become more important. They are no longer just single programs used only by the statistician, but a template that may be used by many programmers. As such, it is important that statisticians also follow GPP.

Although all organizations have their own GPP guideline or SOP, an industry standard guideline is being developed by the PhUSE WG. This will ensure a minimum industry standard that everyone should follow. Asking statisticians to also follow this minimum standard means everyone will benefit, especially the statisticians when they have to validate a program from a programmer, or update a program written by another statistician. The few simple steps in the guideline will help everyone who has to look at a program (in any language).

## Event driven trials in a respiratory setting

**Nick Cowans, Abigail Fuller, Andrew Holmes**

**Veramed**

Event driven trials run until a number of clinical events, typically relevant to the primary end point, have occurred. Such designs can give a study the desired statistical power without having to anticipate true event rates. Subject recruitment may take place over a period of time, with possible geographical temporal variation. However, the study ends for all subjects at a similar point in time, resulting in variable study follow-up for different subjects. Subjects can have less exposure for two reasons: (i) premature treatment withdrawal while the study is ongoing or (ii) recruited later, so unable to continue beyond the study end. While the first reason may potentially be due to adverse events or being from a less healthy subset of the population, the second reason may be more random.

For time to event endpoints followed up post-treatment, this is not a major concern. Censoring means that subjects only contribute to the analyses for the time they were in the trial. However, for

other endpoints, such as repeated measures, care must be taken to ensure that missing data from subjects who withdraw from treatment and subjects who did not have the opportunity to be in the study long enough are treated differently. We present some of the methods used in a large event driven respiratory study to manage this issue.

## Ensuring the quality of your data in Respiratory trials: Data management from a statistical standpoint

**Abby Fuller**

**Veramed**

Large, global late phase studies inevitably involve huge amounts of data of varying quality. Data frequently needs cleaning up prior to locking the database, a responsibility typically lying with data management. The ability to look at multiple extracts of data while the study is ongoing and blinded has enabled us to develop novel methods for increasing the confidence in data quality.

In respiratory, outcomes such as rate of decline of FEV1 can be heavily influenced by outliers. Looking at these in a visual way emphasises the importance of ensuring that these outliers are genuine data points. Similarly, when rates of respiratory tract exacerbations are an endpoint, recording duplicate or overlapping events will alter results. Prior to this work clinicians would spend time looking through vast amounts of data. This talk will present a variety of Patient Profile review tools that has made clinical review a quick and easy process stressing the importance of these data on our endpoints.

## Bayesian Modelling of Disease Progression in Juvenile Dermatomyositis

**1Natacha Gallot, 2Maria De Iorio**

**1Veramed, 2University College London**

This abstract aims to present a Bayesian Modelling of disease progression. The selection of an adequate treatment course or even the development of new treatments, for rare chronic diseases such as Juvenile Dermatomyositis (JDM), is directed by the ability to accurately diagnose the disease and assess its severity at a fixed point in time. In some rare diseases there are as yet no reliable methods or clinical features with which to stratify patients into those at risk of severe complications or to delineate

the rate of disease progression. The purpose of disease progression models is to guide physicians with regard to treatment-related decisions so that patients are brought into remission more rapidly.

To help characterise disease progression over time and to gain a better understanding of the factors influencing disease risk, we propose developing a two state Markov regression model in a Bayesian framework. The transition probabilities between disease and remission state (and vice-versa) are a function of time-homogeneous and time-varying covariates. This latter type of covariate is introduced into the model through a latent health state function which describes subject-specific health over time and accounts for variability among subjects. To highlight clinical variables that have the most effect upon the transition probabilities, variable selection using spike and slab priors was performed. Posterior inference is performed through Markov Chain Monte Carlo methods. The proposed model seems satisfactory enough to describe the disease progression of JDM, but further research on JDM ought to be conducted in order to validate and improve this model.

A case study will be presented to illustrate this approach using data made available from the UK JDM Cohort and Biomarker Study and Repository, hosted at the Institute of Child Health.

## Helping to Drive the Robustness of Preclinical Research

**Katrina Gore**

**Pfizer**

It is hard to pick up a recent copy of Nature, Science or many preclinical biomedical research journals without seeing an article on the issue of non-reproducible research. The pharmaceutical industry is not immune to these issues. Replication of published research findings is a key component of drug target identification and provides confidence to progress internal drug projects. Additionally, we use data from internally developed in vitro and in vivo assays to assess the biological and pharmacokinetic activity, selectivity and safety of novel compounds and make decisions which impact their progression towards nomination for clinical development.

This presentation outlines steps Pfizer is already taking to improve the scientific rigour of experiments through the use of the Assay Capability Tool. The ACT promotes surprisingly basic but absolutely essential experimental design strategies and

represents the distilled experience of the provision of over three decades of statistical support to laboratory scientists. It addresses the age old issue of statistical design, the more recently highlighted issue of bias and the hitherto overlooked issue of whether the assay actually meets the needs of a drug project team.

We believe the Assay Capability Tool is a practical step forward in improving the reproducibility of preclinical research and is central to Pfizer's continued drive to embed excellent statistical design and analysis into all of our research.

## A Simulation study of a controlled imputation approach for analyzing missing data in recurrent events due to early discontinuations

Mattis Gottlow, Sally Hollis, Robert Wan, Ian Hirsch, Annie Darilay, Lisa Weissfeld, Lesley France

AstraZeneca

**Keywords:** Missing Data, Recurrent Events, Clinical Trial Design and Analysis

**Background:** A controlled imputation approach for recurrent events has been developed using a conditional probability relationship between events before and after discontinuation. The treatment effect is often established using an estimand based on the missing at random (MAR) assumption and the jump to reference (J2R) approach is sometimes used to provide a conservative estimate that is not based on the MAR assumption.

**Method:** Simulations were conducted to study the effects of imputation on the estimated treatment effect, its standard error and the power when using the J2R approach for recurrent events with different levels of missing data.

**Results:** We show that when J2R imputation is used, the treatment effect is diluted as expected, and consequently the power is reduced. However the dilution is manageable as long as the number of discontinuations is reasonably low.

**Conclusion:** Our work offers a view of the consequences of using the J2R approach when analyzing missing data in recurrent events due to early discontinuations and serves as a reminder that keeping the amount of missing data low is at least as important as how you deal with it.

## Analysis of Time-Dependent Covariates in a Single Arm Trial

Vincent Haddad

Amgen

Biased analyses comparing responders to non-responders, transplanted to non-transplanted or resected to non-resected subjects are still very common in the clinical literature, posters and oral presentations. This presentation will clarify what are the biases for several common analyses. Alternative methods will be presented: Mantel-Byar test (1974), censored vs. uncensored KM curves (Anderson 1983), Simon & Makuch curves (1984) and Cox model with a time-dependent covariate. However even if these unbiased have no statistical bias; their result interpretations require caution due to confounding factors.

## Considerations for Best Practice for Analysis of Shared Clinical Trial Data

Sally Hollis<sup>1</sup>, Chrissie Fletcher<sup>2</sup>, Frances Lynn<sup>3</sup>, Christoph Gerlinger<sup>4</sup>, Hans-Jörg Urban<sup>5</sup>, Janice Branson<sup>6</sup>, Hans Ulrich Burger<sup>5</sup>

<sup>1</sup>AstraZeneca, <sup>2</sup>Amgen, <sup>3</sup>Biogen Idec, <sup>4</sup>Bayer, <sup>5</sup>Hoffmann-La Roche, <sup>6</sup>Novartis

Increased access to data allows researchers to further explore data collected in previous clinical trials to gain new clinical, scientific and methodological insight. The types of research that might be involved fall into three broad scenarios:

1. To replicate and verify the results in the original study report
2. To investigate the original research questions differently or more thoroughly, including meta-analyses based on IPD
3. To use the data for a research question that is different from the original objective of the trial(s)

Many of the basic principles that apply when planning, executing and interpreting the original analysis of a trial can be applied when additional analyses are conducted or the trial data are re-analysed.

Our aim is to provide guidance to researchers seeking to conduct further analysis of existing clinical trial data on how to:

- assess whether the proposed further exploration of existing clinical data can be supported by the proposed further analyses,
- increase the validity and quality of any further analyses conducted

ducted by ensuring that there is appropriate pre-planning and specification,

- ensure appropriate presentation and interpretation of the results of the further analyses

## Assessment of various continual reassessment method models for dose-escalation phase 1 oncology clinical trials: AZD3514 data and simulation studies

Gareth James<sup>1</sup>, Stefan Symeonides<sup>2</sup>, Jayne Marshall<sup>3</sup>, Julia Young<sup>3</sup>, Glen Clack<sup>3</sup>

<sup>1</sup>Phastar, <sup>2</sup>Edinburgh Cancer Centre, <sup>3</sup>AstraZeneca

**Background:** The continual reassessment method (CRM) is considered more efficient and ethical than standard methods for dose-escalation trials in oncology, but requires an underlying estimate of the dose-toxicity relationship ("prior"). Previously we conducted post-hoc dose-escalation analyses on real-life clinical trial data from an early oncology drug (AZD3514) using the 3+3 method and CRM using six prior approaches; we found each CRM model outperformed the 3+3 method by reducing the number of patients allocated to suboptimal and toxic doses.

**Aim:** To compare the CRM with different prior approaches and the 3+3 method in their ability to determine the true maximum tolerated dose (MTD) of various "true" dose-toxicity relationships.

**Methods:** We will consider seven true dose-toxicity relationships, one based on AZD3514 data and six theoretical with the true MTDs identified as the highest dose where the probability of suffering a DLT is below 33%. For each dose-toxicity relationship we will conduct 1000 simulations and use the 3+3 method and the CRM with six prior approaches to estimate the MTD. This will allow us to understand the effect of the prior and assess performance through the proportion of simulations where the MTD is correct, underestimated or overestimated.

**Results:** Preliminary results have favoured the CRM over the 3 + 3 method. The results of this research will determine the performance of the CRM with each prior approach for dose-escalation clinical trials for various dose-toxicity relationships. As well as showing the potential benefits and pitfalls compared to the 3 + 3 method, we hope that this research will encourage confidence in using the CRM method and identify suitable prior approaches to use.



## Application of assurance calculations and a futility assessment in a Phase III Inflammatory disease trial

Ivana Lazic

GSK

In order to improve the success of clinical trials, the approach to decision making, design and execution of studies has changed. Answers to questions like "What is the probability that our trial will detect a difference between treatments, based on our current belief of the distribution of the true treatment difference?" or "Based on our belief of the possible distribution of the true treatment difference, what is the probability that our trial will achieve the pre-defined success criteria?" have become an essential part of clinical trial development. In order to answer questions like this, assurance described by O' Hagan, J. Stevens and M. Campbell as "the unconditional probability that the trial will yield a positive outcome, where positive outcome means a statistically significant result according to some standard frequentist significance test" will need to be calculated. Assurance, therefore, can contribute to better decision making while designing and reporting a clinical trial. Outlined in this poster is how principles such as conditional power and assurance have been applied to a Phase III inflammatory disease clinical trial.

## Comparing methods for handling missing glycated haemoglobin (HbA1c) values in clinical trials on patients with Type II diabetes

Sophie Lee, Tina Rupnik, Gareth James

Phastar

**Introduction:** In Type II diabetes, glycated haemoglobin (HbA1c) is the most commonly used measure of severity, higher levels are associated with greater mortality and morbidity. Typically in clinical trials, HbA1c will be measured longitudinally, but the level varies considerably between patients and within patients across time. Because of this, simply ignoring missing data or carrying data forward from previous measurements may alter the precision and bias of estimates of the treatment effect. An alternative approach to these methods is multiple imputation, which uses other patient measures to estimate missing values and obtain unbiased and precise estimates if appropriate assumptions are satisfied. The two-fold fully conditional specification (FCS) algorithm imputes

missing values at a given time point conditional on information at the same time point and adjacent time points. A recent study found improved precision in estimates of explanatory variables when using the FCS algorithm compared to multiple imputation on longitudinal data, however it is not known how imputing missing HbA1c data will affect precision and bias of estimates of the treatment effect. We sought to compare methods for handling missing HbA1c data.

**Data:** Anonymised longitudinal primary care data on patients with Type II diabetes from two inner London boroughs between 2007 and 2009.

**Methods:** The results of a recently published manuscript on HbA1c data will be reproduced. We will apply different missingness mechanisms to set data to missing, and explore different methods to handle missing data; complete case analysis, last observation carried forward, multiple imputation and the two-fold FCS algorithm. We will analyse the data using the same methods as the manuscript and compare precision and bias of estimates of the treatment effects.

## Bolstering with Bayes – A framework for interpreting the risk of rare adverse events in the presence of limited clinical trial data

Rachel Moate, Alex Godwood, Jay Zhang

MedImmune

Some classes of drugs are known to be associated with small to modest increases in the risk of a particular rare adverse event. At the end of phase II, drug development teams are primarily interested in the efficacy and emerging safety profile of a new drug. However, there may also be interest in evaluating the risk of an adverse event of special interest. For rare adverse events, limited data are available, and there may be zero occurrences of the event of interest in one or more treatment groups. Making inferences about relative increase in risk using traditional methods thus becomes challenging. In this talk, we will show how utilising historical data for the event of interest by the application of Bayesian methods can provide a framework enabling such interpretations to be made, using informative conjugate priors as described by Kerman<sup>1</sup>. A simulation study evaluating the impact of choice of prior and the performance of decision rules within this framework using R and Rjags will be presented. A case study applying the method to a phase II trial will be described, and the benefits and limitations of the methodology will be discussed.

**Key words:** Bayesian, simulation, rare events

### Reference

1. Kerman, J (2011). Neutral non-informative and informative conjugate beta and gamma prior distributions. Electronic Journal of Statistics.

## Estimation of tolerance limits using a modified Satterthwaite approximation

Thembile Mzolo, Edwin van den Heuvel

University of Goningen and Eindhoven University of Technology

In the pharmaceutical industry, statistical tolerance intervals are commonly used to set specification limits as part of a regulatory necessity for drug substances. Most of the available statistical methods for estimating tolerance limits are model-based specific, with the one-way random effects model being the focal point. To the best of our knowledge, the only approaches applicable to a larger spectrum of random effects model structures are due to Sharma and Mathew (2012) and Hoffman (2010). The former uses the modified likelihood theory, whilst the latter is based on the modified large sample theory. However, the approach of Sharma and Mathew (2012) is computationally intensive and that of Hoffman (2010) is generally conservative. Accordingly, the present study attempts to propose a simple approach for estimating tolerance limits that is applicable to any random effects model. A tolerance factor which depends on the modified Satterthwaite degrees of freedom (van den Heuvel, 2010) is derived. The simulation study showed that, in general the proposed approach gives the coverage which is close to the nominal value. Furthermore, good coverage was observed when a small sample size was considered. One of the main advantages of this approach is that the parameter estimates can be easily obtained using any commercially available statistical software. The current findings add to a growing body of literature on the improvement of the tolerance limits estimation.

### References

- Hoffman, D. (2010), "One-Sided Tolerance Limits for Balanced and Unbalanced Random Effects Models" *Technometrics*, 52, 303-312.
- Sharma, G. and Mathew, T. (2012), "One-Sided and Two-Sided Tolerance Intervals in General Mixed and Random Effects Models Using Small-Sample Asymptotics," *American Statistical Association*, 107, 258-267.
- van den Heuvel, E. (2010), "A Comparison of Estimation Methods on the Coverage Probability of Satterthwaite Confidence Intervals for Assay Precision with Unbalanced Data," *Communications in Statistics - Simulation and Computation*, 39, 777-794.

## 'Hot deck' Imputation: Determining a nonparametric statistical model for the distribution of missing data and its application in a Rate of Decline Analysis

Amy Newlands<sup>1</sup>, Abigail Fuller<sup>2</sup>

<sup>1</sup>GSK, <sup>2</sup>Veramed

Missing data are a potential source of bias when analysing and interpreting the results from study data and unfortunately, all approaches to handling 'missingness' in the analysis rely on assumptions that cannot be verified. There are several existing methods for handling missing data. This poster will show one method of imputation called hotdecking.

In a large outcomes study in approximately 16,000 patients with Chronic Obstructive Pulmonary Disease, one of the secondary endpoints is FEV1 rate of decline. For the main analysis, a random coefficients model is used. However, this approach introduces more weight to subjects with more data; hence a sensitivity analysis of individual regression slopes will be performed. To take account of missing data, and gain understanding of the differences that would have occurred if all subjects had remained on their original treatment and not withdrawn, an imputation approach is used.

In very large studies parametric imputation can become more difficult due to substantial amounts of data. However, in this situation, nonparametric imputation becomes a valid representation of missing data. The missing data is imputed using a double re-sampling approach based on groupings of factors which lead to withdrawal.

This double re-sampling is repeated 100 times and the ob-

served and imputed data combined. An estimate of the treatment difference and standard error is calculated based on Rubin's rule. Since the study has not been un-blinded yet, the imputation approach has been run on blinded data.

## Best Practice for projects involving modelling and simulation

Michael O'Kelly

Quintiles, on behalf of the PSI Special Interest Group for Modelling and Simulation

In 2011 at an EMA-EFPIA workshop Rob Hemmings called for a Best Practice document for projects involving modelling and simulation (M&S), and suggested that PSI might attempt such a document. It was noted at the conference that projects involving M&S can vary greatly in the importance of their contribution to a regulatory submission. The view of EMA was that, depending on the importance of a M&S project, different levels of rigour could apply. The PSI Special Interest Group (SIG) for Modelling and Simulation has drafted a Best practice document that names the elements that need to be addressed for best practice, but allows the project specification to justify its own level of detail and stringency, based on the importance of the project. Key elements that must be addressed include the objectives of the project; the level of pre-specification; assumptions and their justification; the planned analysis and outputs; sensitivity analyses and the level of quality control. At a recent SIG "Hackathon", Professor Chris Jennison described a M&S project, and participants attempted to create a specification for it, using the draft SIG Best Practice. Improved by feedback from the Hackathon, the SIG Best Practice document will soon be submitted to PSI for review.

## Rank-based estimation for the non-normal general linear model: a tool for the industry

John Pemberton

Phastar

Non-normal errors in the linear model have long been problematic. Either a log transform is used where one simply hopes for the best, or a method based on the Rank Transform (1) adapted from (2) is applied to test for and estimate a treatment effect. Neither of these methods deals with the issue adequately and the rank transform certainly cannot cope at all when there are interactions. A superior alternative is Rank-based estimation (based on ranks

of residuals) whose theory and application has been developed since about 1970. For simplicity we restrict attention to the linear model, but the methods extend to other models including survival, mixed effects and so on. The publication of a number of books (3; 4), an excellent R package and recent papers providing applications to the clinical trial area (5; 6) make this approach easily accessible. It far outperforms the Rank Transform approach, which neither competes in quality of inference nor range of application.

We provide a brief introduction to the method and illustrate its superiority over both least squares and rank transform methods using both real and non-normal simulated data.

### Bibliography

1. Stokes, M.E., Davis, S.D., Koch, G.G. *Categorical Data Analysis Using SAS®*, Third Edition. Cary : SAS Institute, 2012. 978-1-60764-664-8.
2. Quade, D.E. *Rank Analysis of Covariance*. J. Amer. Stat. Ass. 1967, Vol. 62, 320.
3. Hettmansperger, Thomas P and McKean, Joseph W. *Robust Nonparametric Statistical Methods*. Boca Raton : Chapman & Hall, 2011.
4. Kloeke, J and McKean, J W. *Nonparametric Statistical Methods using R*. Boca Raton : Chapman & Hall, 2014.
5. Kloeke, J D and Cook, T. *Nonparametric covariate-adjusted hypothesis tests using R estimation for clinical trials*. 2015 (to appear).
6. Rashid, M M, McKean, J W and Kloeke, J D. *R estimates and associated inferences for mixed models with covariates in a multi-center clinical trial*. *Statistics in Biopharmaceutical Research*. 2012, Vol. 4.

## Investigating the performance of a Poisson regression model, negative binomial, zero-inflated Poisson and zero-inflated negative binomial regression model in presence of over-dispersed counts

Anna Rigazio, Audrone Aksomaityte

Phastar

In asthma clinical trials Poisson regression is frequently used to analyse exacerbation rates, assuming that the mean occurrence rate of the event is equal to its variance. Asthma exacerbation



data are often characterised by over-dispersion and frequent zero-count observations. Thus, a Poisson regression might fit these data poorly and other generalised linear models could perform better. When the variance is higher than the mean event rate, a negative binomial (NB) regression model should be preferable. Zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) models are also used to avoid the underestimation of rates of excess zero-count events.

We will investigate how the performance of a Poisson regression model, as well as a NB, a ZIP and a ZINB regression model is affected by the following parameters:

- sample size;
- average event rate;
- over-dispersion parameter;
- number of individual with zero exacerbations.

We will simulate asthma exacerbation data in order to identify potential thresholds for these parameters to use as guidance in the choice of the best fitting model.

## Statistical models for de facto estimands - beyond sensitivity analysis

James Roger

LSHTM

Regulators throughout the world are moving the emphasis for estimands in confirmatory studies away from efficacy and towards effectiveness. There is the prospect that de facto rather than de jure estimands will soon be required for primary analyses. If so, what will replace MMRM as the default approach for handling early withdrawal in longitudinal studies?

Currently such de facto estimands lie in the domain of sensitivity analysis using multiple imputation, often known as reference-based imputation. In such sensitivity analyses we modify the data generating model but retain the analysis model in its original form. But for a primary analysis in a confirmatory trial we need to start again, with a model where the data generation and the analysis models are congenial.

Here we describe a joint modelling approach for withdrawal and outcome with distinct but correlated models for outcome before and after withdrawal based on extrapolation factorization rather than the more usual selection or pattern-mixture approaches. The model can be fitted using maximum likelihood or in a Bayesian way. Indeed the inclusion of parameters which are not estimable

from the data is well suited to a Bayesian approach. The equivalent of least-squares means and their differences are accommodated by evaluating the expected response under estimated parameters (either MLE or based on a sample from Bayesian posterior) and a fixed set of covariate values.

These estimands are marginal predicted values involving parameters of both the withdrawal and outcome processes. As such this requires integration. For some estimands it requires quite complex numerical integration, while for others algebraic solutions are possible.

The main message of the talk is how moving towards de facto estimands for primary analysis in confirmatory trials will require both careful thought and most likely the development of new computational tools.

## Practical considerations in fitting generalised Gamma distributions for HTA

Stuart Spencer, Richard Lawson

AstraZeneca

For HTA assessments multiple distributions must be considered when modelling time to event endpoints. The guidance suggests using AIC to identify the best fitting model. Common software packages such as R and SAS offer many different optimisation algorithms for parameter and likelihood estimation and these can give different AIC values therefore when following the guidance consideration should be given to the optimisation method.

Among the distributions recommended for consideration by the guidance is the Generalised Gamma distribution. We have noted that, when using small datasets, the Gamma may be the best fitting according to the AIC but then give extreme parameter estimates which are unsuitable for extrapolation.

Four percent of health technology appraisals at NICE (UK) use the Generalised Gamma distribution to generate estimates of survival. Using two AstraZeneca datasets, in different therapy areas; from products in development; this research describes problems in fitting a Gamma function in time to event analyses and considers whether implementing different algorithms in R is pragmatic.

## Increasing the Efficiency of Early Phase Decision Making studies by using a continuous endpoint within a Bayesian Framework

Foteini Strimenopoulou<sup>1</sup>, Emma Jones<sup>2</sup>, Ros Walley<sup>1</sup>  
1UCB, 2Veramed

Typically for any disease, there will be a gold standard efficacy measure. Whilst such a measure may well have considerable discriminant ability, be widely accepted and easy to interpret, it may also have significant drawbacks. For example, it may be a binary measure and therefore require a large sample size to detect clinically significant differences. Secondly, it may not reflect small improvements for an individual, which if seen over a short time period, could correspond to more marked differences in a longer study. For early phase studies using a related continuous endpoint within a Bayesian framework may be very beneficial.

In this talk we give an example from the inflammation therapeutic area at UCB, in particular from rheumatoid arthritis. The binary endpoints ACR20, ACR50 and ACR70 are often used to assess efficacy in RA. However, these endpoints are derived from a continuous endpoint named ACRn. Here, we will illustrate the use of the continuous measure, ACRn, under an appropriate transformation, as the preferred measure for early decision making. We will show the relationship between this continuous endpoint and the binary ones, based on their definitions and through consideration of historical data. Finally, we will discuss the resource savings that can be made.

## An Investigation into Overfitting

James Sykes, Omar Fathi

Phastar

Statistical modelling techniques such as linear regression and repeated measures are often used in pharmaceutical research in order to understand the effects of a pharmaceutical product. When these models are fitted with multiple covariates and various complexities, such as functions of covariates, overfitting may not always be obvious. We describe overfitting here as the unneeded complexity of a model due to the use of too many predictors in relation to the number of observations and the unneeded use of functions of these predictors. Thus, as a predictive tool for subsequent data, an overfitted model is ineffective and yields

spurious results and therefore creates uncertainty regarding the scientific value of the findings that are ascertained. This poster aims to shed new light over the topic via the application of a simulation study investigating different models varying in complexity by observing model fitting statistics. We intend to investigate the scenarios when the data is replicated and the model is used on future (replicated) data, incorporating repeated measurements using SAS Proc Mixed. This will then allow us to compare the relationship between the complexities of a model and the degree of overfitting observed.

## Key considerations for fitting Logistic Regression in SAS®

Lyn Taylor, Helen Brown

PAREXEL

With advancements in statistical software, it is simple to fit statistical models with little statistical knowledge. Incorrect model specification and a lack of model assumption checking can easily result in invalid analyses. As statisticians we need to ensure we fully understand the analysis being performed behind the software coding being used. This poster identifies potential issues which should be considered before interpreting the results from a PROC LOGISTIC in SAS®. A review of how to use the design matrix to create contrast statements is presented for individual and pooled treatment comparisons. Different methods of model parameterisation are discussed along with details of how to obtain estimates of the treatment differences and confidence intervals. The importance of model convergence, ensuring the correct outcome is being modelled and the dangers of fitting too many factors are also discussed.

## First Experience in Observational Research – A Statistician's Perspective

Chris Toffis

Amgen

The objective of this poster is to describe my initial experiences of working as a statistician on observational research studies. I introduce the common types of observational study designs and discuss the features of prospective and retrospective studies. Some challenges that have arisen during the conduct of our observational studies are then presented, such as the presence

of historic contradictory data with potentially no means of querying to obtain resolution. Lastly, I outline the potential biases inherent in observational research and discuss some of the statistical approaches that will be applied to overcome the limitations in our studies.

## Mediation analyses for trials of parenting programmes with missing values in baseline and parenting measures

Angela Cheng Zhang<sup>1</sup>, Professor Stephen Scott<sup>2</sup>,  
Professor Sabine Landau<sup>2</sup>

<sup>1</sup>Novartis; <sup>2</sup>King's College London

Parenting programmes are the most effective intervention to change persistent child anti-social behaviour and widely used, but little is known about the mechanisms through which they work and hence how to improve them. The theoretical model underlying parenting programmes assumes that child outcome can be improved by interventions that improve parenting. Trials of parenting interventions routinely evaluate the effectiveness of the treatment in terms of the clinical outcome (child anti-social behaviour) and putative mediators (parenting practices). However, they tend not to carry out formal analyses to explicitly decompose total treatment effects into indirect (mediated) and direct (non-mediated) components. In this project we will use data from three randomised control trials of the Incredible Years (IY) parenting training programme to assess the mediation processes through which these interventions work. Practically, parenting behaviour is a difficult construct to measure and studies typically employ multiple measurement methods.

Traditional mediation approach such as the regression approach by Baron and Kenny[1] assumes no unobserved confounding of the effect of the mediator on the outcome. This assumption leads to a need of including all measured pre-randomisation and post-randomisation confounding variables in the B&K mediation model. Our datasets are subject to missing data: (1) Roughly a third of the families have incomplete baseline characteristics variables. (2) Missing values are present in both parenting practices variables and child outcomes of interest. We apply multiple imputation (MI)[2] to traditional mediation analyses so that analyses are valid under a realistic missing at random assumption. Furthermore, including auxiliary multi-informant

parenting behaviour variables in the imputation model allows us to exploit all the available information provided in the trials. The implementation of the approach resolves the following practical and technical challenges: (i) It is inapplicable to include all the measured potential confounders in the model due to small sample size. We proposed an approach to select confounding variables for inclusion as covariates in the regression equations. (ii) IY parenting programme is a group training therapy and the therapy groups are only applicable in the active intervention arm. The MI step allows for clustered data that only explain variability in the active intervention arm. The estimate of causal mediation parameters is calculated by averaging the estimates obtained from each imputed dataset. Nonparametric bootstrap approach is developed for drawing statistical inferences to assess relevant mediation parameters with MI applied for each bootstrap dataset with missing values. Two mediators, parental criticism and parental warmth, are detected to mediate 38.57% and 22.03% of the total effect respectively.

Instrumental Variables (IV)[3] methods have more recently been advocated for addressing causal questions in social and public health. In this project, we further developed the IV approach to handle confounding issues of mediation analyses in the parenting programme context. After setting up a list of inclusion criteria, we selected the interaction terms between randomization and baseline parental characteristics, the interaction terms between randomization and treatment process variables as IVs to evaluate the mediation effects of IY parenting programmes. In combination with MI, our approach is robust for dealing with unmeasured confounding issue in parenting mediation analyses with missing data. The IV analyses results support the same two mediators with reduced mediation effects estimate and wider confidence interval.

1. Baron, R.M. and D.A. Kenny, The moderator mediator variable distinction in social psychological-research - conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 1986. 51(6): p. 1173-1182.
2. White, I.R., P. Royston, and A.M. Wood, Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, 2011. 30(4): p. 377-399.
3. Wooldridge, J.M., *Econometric Analysis of Cross Section and Panel Data* 2002, Cambridge: Massachusetts Institute of Technology.



## Assessing the cardiovascular risk of anti-diabetic therapies in patients with type 2 diabetes mellitus

Richard C. Zink

JMP Life Sciences, SAS Institute

Recent guidance from the United States Food and Drug Administration (US FDA) and European Medicines Agency present recommendations to assess cardiovascular (CV) safety for non-insulin anti-diabetic therapies in patients with type 2 diabetes mellitus (T2DM). In particular, the risk of major adverse CV events, which includes CV death, non-fatal myocardial infarction and non-fatal stroke events, is assessed in two stages in the US FDA guidance. Stage 1 is a pre-market evaluation of the novel compound to placebo testing whether the upper bound of the 95% confidence interval of the hazard ratio is  $< 1.8$ . Assuming the drug application is otherwise acceptable, if the CV criteria is met the sponsor obtains full marketing approval for the new drug. In Stage 2, the sponsor must evaluate the post-market criteria testing the hazard ratio against a more stringent upper bound of 1.3. This approach is to strike a balance between providing evidence on cardiovascular safety to reassure patients and excessive delay of novel therapies reaching the marketplace. To understand the impact of FDA guidance on T2DM development programs, we reviewed drug applications for treatments approved by the US FDA during 2002-2014. In this talk, we summarize the CV assessment strategies applied in practice, and describe the advantages and disadvantages of individual methods. The implications of the above regulatory framework, particularly in regards to the size of the safety database and the confidentiality of interim results, are discussed. This work is presented on behalf of the Safety Working Group of the Biopharmaceutical Section of the American Statistical Association.

identify patients receiving additional benefit from the treatment in terms of an efficacy response. Alternatively, finding subgroups based on important safety endpoints could be considered to determine those individuals experiencing a reduced risk of key adverse events, or to identify subjects for whom the new therapy may be inappropriate. Tree-based methods are naturally compelling in this context, and we review a few popular approaches that leverage recursive partitioning and hierarchical clustering. These analyses can be interpreted as finding the right patients for a given treatment. We compare them to optimal treatment regimes, which alternatively focus on finding the best treatment assignment (drug and/or dose) for each patient.

## Subgroup analyses for personalized medicine

Richard C. Zink, Russell D. Wolfinger

JMP Life Sciences, SAS Institute

In contrast to the “one-size-fits-all” approach of traditional drug development, the need to locate subjects with an enhanced treatment effect is a critical component for modern tailored therapeutics or personalized medicine. Typically, the goal is to

# PSI Conference 2015: Collated Abstracts

## Tuesday:

### Estimands and Sensitivity Analyses

Defining the primary objective of a clinical trial in the presence of non-compliance or non-adherence to the assigned treatment is crucial for the choice of design, the statistical analysis and the interpretation of the results.

At first glance this seems obvious, however, primary objectives stated in clinical trial protocols often fail to give a precise definition of the measure of intervention effect. The impact of potential confounding, e.g. due to non-compliance, missing data, treatment switching / discontinuation or intake of rescue medication, is frequently not taken into account when defining the intervention effect of interest.

The need for a structured framework to specify the primary estimand (i.e. 'what is to be estimated') was highlighted in the context of missing data in the National Academy of Science document "The Prevention and Treatment of Missing Data in Clinical Trials" (2010). However, the need for clearly defined estimands applies to a broader setting.

In these two sessions we will discuss the need for this framework, the definition of estimands, the choice of estimands in different settings and the role of sensitivity analyses. These aspects will be discussed from a regulatory, industry and academic point of view.

#### Chair

Lesley France (AstraZeneca)

#### Speakers:

1. Rob Hemmings (MHRA) will motivate the need for the estimand concept
2. Tom Permutt (FDA) will motivate the new framework further from an US perspective
3. Chrissie Fletcher (Amgen) will give the industry perspective as the EFPIA representative in the ICH working group
4. James Carpenter (LSHTM) will give an academic perspective on the topic
5. Alan Phillips (ICON) will give a PSI perspective + serve as an introduction to the panel discussion

### Estimands and Sensitivity Analyses Panel Discussion

Following a short presentation from Alan Phillips all five speakers and James Roger (LiveData) will take part in a panel discussion so have your questions at the ready.

### Adaptive Designs: Reflections on Their Current Use in Drug Development

In this first of two sessions on adaptive designs, speakers will consider the current landscape in the use of adaptive designs. This will draw on experiences from the regulatory perspective, research into attitudes from both the private and public sector points of view and some areas of current interest in academia.

#### 1. Rob Hemmings (MHRA): "Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency"

Since the first methodological publications on adaptive study design approaches in the 1990s, the application of these approaches in drug development has raised increasing interest among academia, industry and regulators. The European Medicines Agency (EMA) as well as the Food and Drug Administration (FDA) have published guidance documents addressing the potentials and limitations of adaptive designs in the regulatory context. Since there is limited experience in the implementation and interpretation of adaptive clinical trials, early interaction with regulators is recommended. The EMA offers such interactions through scientific advice and protocol assistance procedures.

We performed a text search of scientific advice letters issued between 1 January 2007 and 8 May 2012 that contained relevant key terms. Letters containing questions related to adaptive clinical trials in phases II or III were selected for further analysis. From the selected letters, important characteristics of the proposed design and its context in the drug development program, as well as the responses of the Committee for Human Medicinal Products (CHMP)/Scientific Advice Working Party (SAWP), were extracted and categorized. For 41 more recent procedures (1 January

2009 to 8 May 2012), additional details of the trial design and the CHMP/SAWP responses were assessed.

A summary of the characteristics of the submitted studies along with the subsequent advice will be presented. In addition, case studies are presented as examples.

Joint work with Amelie Elsässer, Jan Regnstrom, Thorsten Vetter, Franz Koenig, Martina Greco, Marisa Papaluca-Amati and Martin Posch.

#### 2. Munyaradzi Dimairo (University of Sheffield): "Bridging the gap in take-up of adaptive designs in confirmatory trials: results from interviews and surveys of key stakeholders in trials research"

Authors: Munyaradzi Dimairo, Steven A Julious, Susan C Todd, Jon P Nicholl

Routine use of adaptive designs, especially in the confirmatory phase of drug development has been lagging behind attention given to this topic in the literature. Building on previous related work, we conducted qualitative interviews of key stakeholders (predominantly UK public sector) in trials research to explore barriers to implementation of these methods and opportunities for greater use in the future. We further undertook follow-up quantitative surveys both in the private and public sector to generalise the findings and to compare and contrast the two sector perspectives. In this talk results and findings from this work will be described. Most importantly, we rank priority areas and suggest potential solutions to overcome some of the obstacles to facilitate successful implementation of adaptive designs in a confirmatory setting in the future.

#### 3. Sue Todd (University of Reading): "Methods of analysis: Catching up with developments in design?"

Considerable literature exists on methodology for adaptive designs, primarily on the topic of designing an adaptive study, focusing often on the control of type I error rate. Less research has been undertaken into inference following the completion of an adaptive study. This talk will consider some of the available methods of analysis following an adaptive design, highlighting issues of bias and accuracy. The use of such methods in current trials will be explored.



## Risk-Based Monitoring

Risk based monitoring (RBM) has been growing over the last few years, as we look for ways to reduce the cost of clinical trials, whilst ensuring the data quality does not drop. Today's presentations will look into different applications of RBM, uncovering some of the complexities and learnings that has been encountered as we gain more experience. We look forward to seeing you there!

### Speakers:

#### 1. Richard Zink (JMP Life Sciences, SAS Institute): "Analytical Considerations for Risk-Based Monitoring"

Central computerized review of clinical trial data enables risk-based monitoring (RBM) to determine if sites should receive more extensive quality review or intervention. The availability of extensive logic and validation checks to detect outliers and implausible values early in the clinical trial not only ensures data quality, but can be used to identify instances of data fabrication and other forms of misconduct. This presentation discusses analytical considerations for RBM, including supervised and unsupervised methodologies, and the need to consider both sets of approaches in practice. Regulatory guidance and the TransCelerate position paper on RBM methodology motivate the discussion

#### 2. Shafi Chowdhury (Shafi Consultancy): "Risk-based Monitoring – The Score Card Approach"

Risk-based monitoring has been sweeping across the industry over the past few years, and its aims and implications are not always clear. There are discussions about whether this relates to how risky the approach is, or if it is the risk of getting bad quality data. In reality, it is the coming together of a knowledge based approach towards improving the quality of data. Reviewing key data points in a manual and or automated process and to dictate where limited resources should be targeted.

Centralised risk-based monitoring is a better term for this approach. All risks are identified and steps are defined to mitigate them in advance in a Risk Management Plan. Programs are used to check the data quality of each site, including fraud detection, and combining this with knowledge from CRAs performing on-site visits, it is possible to determine the relative risk of each site. This paper will look at one method of calculating risk and what actions can be taken based on the risk.

#### 3. Alun Bedding (Roche): "The Use of Statistical Methods in Risk Based Monitoring"

Authors: Alun Bedding, Chris Wells - Roche

Risk based monitoring has been identified by the FDA as a way to improve data quality and ensure patient safety. The use of statistical methods enables a sponsor to look for non-systematic patterns in the data that cannot be picked up using standard tests. These patterns in the data can cause issues with the data and may be indicative of data misconduct. Many authors have addressed this issue and vendors are now proposing software with which to perform these analyses. The impact of findings has reaching impact in the integrity of a trial and maybe a submission. This presentation will outline some of the main methods given by authors and will illustrate using JMP®/Clinical software. Real clinical trial data will be used, where possible to illustrate the methods.

## Contributed Papers Session: Modelling and Simulation

The following speakers have been selected from the contributed abstracts received to talk on modelling and simulated related topics.

### Speakers:

#### 1. Sinead Hamilton (Quintiles): "Using ideas of Best Practice for Modelling and Simulation in a project to simulate a Negative Binomially Distributed Recurrent Event Dataset"

Authors: Sinéad C. Hamilton, Michael O'Kelly (Quintiles)

In 2011, Rob Hemmings (MHRA) called for a Best Practice document for projects involving Modelling and Simulation. His call was supported widely by industry practitioners. In March 2015, PSI's Modelling and Simulation Special Interest Group (SIG) ran a workshop to finalise such a Best Practice document, using a draft proposed by the SIG. A number of example projects have used the SIG's draft Best Practice document. This presentation shows how Best Practice could be followed, using the SIG's Best Practice document, where the project involved simulating outcomes that follow the Negative Binomial distribution.

The Negative Binomial distribution can be regarded as a Poisson distribution with an effective intensity modified multiplicatively by a gamma-distributed random variable. The presentation shows how Negative Binomial outcomes can be simulated in a simple manner by using the asymptotic equivalence of the binary and the Poisson distributions.

A feature of the project was that we encountered problems with

the specification of simulations when it came to the detail of achieving one of the objectives. This resulted in multiple revisions of the specification. We describe this process in our presentation. The finalisation of the specification is described; and the results are presented, with a description of how the conclusions help to answer the questions posed in the objectives specified for this Modelling and Simulation project.

#### 2. Gautier Paux (IRIS): "Key Principles of Clinical Trial Simulations to Improve the Probability of Success in Late-Stage Trials"

Authors: Gautier Paux (IRIS), Alex Dmitrienko (Quintiles)

Confronted with the increasing cost, duration and failure rate of new drug development programs, the use of innovative trial designs and analysis strategies has considerably increased over the past decade. In this context, clinical trial simulations take a crucial and invaluable role to support a thorough assessment of the operating characteristics and performance of candidate designs and strategies. Before conducting a trial, simulation-based methods allow clinical trial sponsors to evaluate the effect of individual design or analysis parameters (as well as their synergic effect) on relevant criteria of trial success. Additionally, they facilitate the assessment of risks and benefits associated with each candidate design and analysis strategy and provide justification of parameter choices. Recently, the Mediana R package has been developed to provide a standardized approach to clinical trial simulations to facilitate a systematic simulation-based assessment of trial designs and analysis methods in clinical trials or across development programs. This package supports a broad range of trial designs and analysis methods typically used in late-stage trials. In this presentation we will discuss key principles of clinical trial simulations in the context of Phase II and Phase III trials to arrive at the optimal selection of design and analysis parameters.

#### 3. Euan Macpherson (AstraZeneca): "Design of clinical trials in the presence of a delayed treatment effect"

Authors: Euan Macpherson, Mary Jenner, Paul Metcalfe (AstraZeneca)

Background: Power and sample size calculations for clinical studies with time to event endpoints are routinely produced assuming a constant treatment effect, implying only the number of events and not overall study maturity need be considered to achieve statistical significance with a given power. However, some targeted therapies (e.g. in Immuno-Oncology), may have a mode

of action resulting in a delay before a treatment effect emerges on the survival curves. Analysis with more weight given to earlier events would have lower probability (power) to detect a real treatment effect where this emerges at later maturity.

**Aims:** Illustrate the implications for trial operating characteristics where a delayed treatment effect is anticipated at the trial design stage.

**Methods:** Statistical software has been developed (R package) to enable study teams to investigate delayed treatment effect scenarios and calculate required number of events and power.

**Results and Discussion:** Trial design scenarios are explored with the software and the output is interpreted in comparison with routine calculations to highlight the risk of under-powering a study where there is potential for a delayed effect. Consideration of trial maturity is strongly recommended in such situations to ensure adequate power.

## How Can Data Sharing Help You: Real life Examples

Though the more formal process for sharing data are relatively new. Companies and organisations have been sharing data for a long time to help answer scientific questions and improve our joint knowledge in diseases including new endpoint definitions. This session takes a practical approach to three different therapeutic areas and how data sharing has been applied to move research forward and help patients.

### Speakers:

#### 1. Brian Tom (MRC): “The RA-MAP Experience: Investigating Predictors of Remission by Combining Data from Trials in Rheumatoid Arthritis”

In this talk I describe the experience of industry and academic experts working together as part of the MRC/ABPI Inflammation and Immunology Initiative's RA-MAP consortium in Rheumatoid Arthritis (RA). In particular, I focus on the challenges, opportunities, outcomes and lessons learnt when combining the controlled arm data from randomised controlled trials of biologics in RA for investigating predictors of remission.

#### 2. Dr Paul Wren (GSK): “Progress in Public Private Partnerships engaged in the Dementia Challenge”

Working through partnership across public private sectors enables the sharing of ideas, data and resources to provide open science

platforms for advances in Medical Research. The recently initiated Dementias Platform UK will be used as one example of how academia and industry are working together to increase the understanding of the Dementias to accelerate the delivery of medicines to patients. A holistic view of the early development of the platform will be shared from an Industry perspective with a specific focus on the establishment of core infrastructure and methodologies to enable secure data collation and analysis from across multiple large diverse clinical cohorts across the spectrum of Dementias. Using a range of methodologies including neuroimaging, genetics, fluid and physiological biomarkers, cognitive testing and clinical outcomes, innovative experimental hypothesis testing and collation of longitudinal data will help to enhance our knowledge and probabilities to ultimately deliver disease modifying medicines.

#### 3. Liz Zhou (Project Datasphere): “Unleash the Potential of Clinical Trial Data: the Project Data Sphere Initiative”

Researchers are working tirelessly and new advances are constantly being discovered, yet every day, tens of thousands of our loved ones lose their battle with cancer. Sadly, we're losing nearly the same number of people today as we were 40 years ago.

Meanwhile, huge amounts of clinical trial data are sitting within repositories of commercial and public databases collecting dust because they are typically used just for a single purpose. What if we could share our collective historical cancer research data in a single location? The Project Data Sphere initiative was conceived from this original idea; it has since had a successful launch in April 2014 and is well on its way to reach 25,000 patient lives by one year anniversary.

A goal of the Project Data Sphere initiative is to spark innovation through access to comparator arm data from historical cancer clinical trials. The data can allow for more efficient research through improved trial design, reduced duplication, as well as the development of broader data standards. A platform across all cancer types, open to all researchers, may unleash the full potential of the data to advance research and benefit cancer patients.

The true power of this platform will come from an increasing volume of data and the continuing engagement of a global community focused on finding solutions for cancer patients. Imagine what will happen when the entire cancer community joins efforts.

## Adaptive Designs: Case-Studies of Recent Experiences in Implementation

In this second of two sessions on adaptive designs, the focus will be on implementation. Speakers will draw on their experiences of using adaptive designs in practice, specifically highlighting advantages and disadvantages / what went well and what did not!

### Speakers:

#### 1. Stephane Heritier (Monash University): “A single pivotal adaptive trial in infants with proliferating hemangioma: rationale, design challenges, experience and recommendations”

This work reflects on our experience when designing and analysing an adaptive confirmatory trial (previously referred to as a seamless Phase II/III trial) in infants with hemangioma over the 2009-2013 period. At the end of the first stage (Phase II) an interim analysis was conducted by an independent data monitoring committee allowing three possible adaptations: 1) selection of one or two active treatment regimens for further study in the second stage (Phase III); 2) sample size reestimation; 3) early stopping for futility. The trial design was defended before the FDA and the EMA prior to trial initiation in 2010, and the primary endpoint was analysed in 2012. Marketing authorisation for the pediatric drug Hemangeol (propranolol hydrochloride) was granted to the sponsor, Pierre Fabre Dermatologie, by the FDA (for orphan indication) in March 2014, and by the EMA in April 2014 under different spelling Hemangirol. Propranolol hydrochloride is the first and only approved treatment for “proliferating infantile hemangioma requiring systemic therapy”. This single pivotal trial is one of the first adaptive confirmatory trials to be conducted successfully in the regulatory setting.

Joint work with Caroline Morgan (Cytel), Serigne Lo (Sydney University) and Jean-Jacques Voisard (Pierre Fabre Laboratories)

#### 2. Marc Vandemeulebroecke (Novartis): “To seamless or not to seamless? Lessons learned from four case studies”

**Background:** Inferentially seamless studies are one of the best known adaptive trial designs. Statistical inference for these studies is a well studied problem. Regulatory guidance suggests that statistical issues associated with study conduct are not as well understood. Some of these issues are caused by the need for early pre-specification of the phase III design and the absence of sponsor access to unblinded data. Before statisticians decide to choose a seamless IIb/III design for their programme, they should consider whether these pitfalls will be an issue for their programme.



**Methods:** We consider four case studies from different pharmaceutical sponsors. Each design met with varying degrees of success. We explore the reasons for this variation to identify characteristics of drug development programmes that lend themselves well to inferentially seamless trials and other characteristics that warn of difficulties.

**Results:** Seamless studies require increased upfront investment and planning to enable the phase III design to be specified at the outset of phase II. Pivotal, inferentially seamless studies are unlikely to allow meaningful sponsor access to unblinded data before study completion. This limits a sponsor's ability to reflect new information in the phase III portion.

**Conclusions:** When few clinical data have been gathered about a drug, phase II data will answer many unresolved questions. Committing to phase III plans and study designs before phase II begins introduces extra risk to drug development. However, seamless pivotal studies may be an attractive option when the clinical setting and development programme allow, for example, when revisiting dose selection.

**References:** Cuffe, Lawrence, Stone, Vandemeulebroecke: "When is a seamless study desirable? Case studies from different pharmaceutical sponsors." *Pharmaceutical Statistics*, to appear 2014  
Joint work with Robert L Cuffe (ViV Healthcare), David Lawrence (Novartis) and Andrew Stone (AstraZeneca).

### 3. Kirsty Hicks (GSK): "An adaptive dose ranging Phase IIb study in patients with Systemic Lupus Erythematosus: An experience from the beginning to the end"

Systemic lupus erythematosus (SLE) is a chronic autoimmune disorder characterised by autoantibody production and abnormal B lymphocyte function. This disease is more common in women (approximately 90% of patients) than men and prevalence varies with race. Systemic lupus erythematosus can lead to arthritis, kidney failure, heart and lung inflammation, and central nervous system changes. The presentation will outline how an adaptive phase IIb study in patients with this disease was designed, logistically run and then analysed. The main objective of the study was to investigate the dose response relationship across four active doses and placebo with a key pharmacodynamic marker initially, followed by the primary efficacy endpoint (SLEDAI) at later decision points. The study was adaptive in nature where a number of interim analyses were incorporated to include various options to drop doses, stop the study for futility, stop the study for safety and even change the characteristics of the patient population.

### 4. Thomas Zwingers (CROS NT): "Murphy's law in Adaptive Designs – what can go wrong, will go wrong"

The primary goal of adaptive designs in drug development is to shorten timelines and minimize the risk of making uninformed and incorrect decisions. A common feature that all adaptive designs share is that they summarize the information in a clinical trial at a very early stage, usually during interim analyses.

Study protocols which foresee adaptive designs require thorough planning, but the best planning can be overruled by reality. The most sensitive areas with respect to the planning are patient recruitment, patient's baseline characteristics and selection of hypotheses. Usually the interim analysis is planned at a certain fraction of the calculated total sample size. A higher recruitment rate as anticipated will cause problems with respect to an effect called "overrunning". In a combinational adaptive design the comparability of the patient cohorts is an essential pre-requisite for the global interpretation of the tested hypotheses. If the cohorts differ too much, the interpretation of the hypothesis might be questionable. Especially in dose-finding studies, testing hypotheses in a hierarchical way is common praxis to minimize the sample size. But very often the dose-response curve is not increasing over all dosages.

We will show examples on the above mentioned problems, which caused serious problems for the studies.

### Break-out Session: Challenging Study Designs

Grappling with a tricky trial? You need to pick the brains of a room full of statistical experts at our "Challenging Study Design" breakout session!

Following the success of last year's conference break-out sessions, we are again running two round table discussion forums. The assembled audience for this session will be divided into groups, with each group being given one or more brief scenarios of studies that are deemed challenging to design in some way. After a period of time for an informal discussion of the key issues and possible designs, the full audience reconvenes to hear the views of all groups. This session is open to everyone no matter your experience. The focus is on interaction, idea sharing and discussion with peers and is not lecture based.

### Data Transparency: - To Infinity and Beyond

As we all become more aware of both what our responsibilities are in sharing our trial data, as well as the opportunities it can bring us, this session aims to draw on different perspectives, to share where we are now and what the future could be and includes an overview of the new EMA guidance on Data Transparency. This session will contain four short presentations from our speakers, Francesco Pignatti (EMA), Frank Langer (Eli Lilly), Trish Groves (BMJ), Sarah Nolan (Uni. of Liverpool) followed by a panel discussion. Come prepared for some interesting discussions and your questions!

#### Speakers:

#### 1. Francesco Pignatti (EMA): "Implementing the new EMA policy on publication of clinical data"

The European Medicines Agency has in recent years made many efforts to further improve its transparency. In this context, the creation of a policy on pro-active publication has been a significant step forward. On 1 January 2015 the new EMA policy on publication of clinical data for medicinal products for human use entered into force. Under this policy, the Agency proactively publishes the clinical reports submitted as part of marketing-authorisation applications for human medicines. This represents the first phase of the policy, concerning overviews and clinical study reports with some appendices, and excludes independent patient data (IPD). The talk will focus on the new EMA policy including European legislative and regulatory framework and anonymising data from study reports.

#### 2. Frank Langer (Eli Lilly): "Disclosure of clinical trial data - Challenges and Opportunities for Statistics"

Transparency of clinical research has been an evolving topic over several years (1). The discussions have included the sharing of summary data and more recently sharing individual patient data from clinical trials (2;3). Statisticians play a key role in helping to balance the different dimensions of responsible clinical data sharing by generating more useful scientific insight while safeguarding patient privacy. In this context opportunities to contribute to the scientific dialogue and foster robust research will be discussed (4;5).



1. EMA Policy/0070 European Medicines Agency policy on publication of clinical data for medicinal products for human use. [http://www.ema.europa.eu/docs/en\\_GB/document\\_library/Other/2014/10/WC500174796.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/Other/2014/10/WC500174796.pdf) 2014-10-07

2. ClinicalStudyDataRequest.com

3. Responsible data sharing. Available at: EFPIA website: <http://transparency.efpia.eu/responsible-data-sharing>

4. Fletcher C, Driessen S, Burger HU, Gerlinger C, Biesheuvel E; EFSPI.

European Federation of Statisticians in the Pharmaceutical Industry's position on access to clinical trial data. Pharm Stat. 2013; 12(6): 333-6.

5. Jeffrey M. Drazen, M.D. Sharing Individual Patient Data from Clinical Trials. N Engl J Med 2015; 372:201-202 January 15, 2015. DOI: 10.1056/NEJMp1415160

Project and the [clinicalstudydatarequest.com](http://clinicalstudydatarequest.com) platform?

Finally, this presentation will include an update on The BMJ's policy on data sharing on request for all trials of drugs or devices.

### 3. Sarah Nolan (Uni. of Liverpool): "Data Transparency – an academic's voyage"

I will describe the activities of an academic researcher in the context of clinical trial data analysis including the importance of data access and data transparency in evidence based medicine. I will discuss essential requirements for researchers in the data transfer process and the obligations of the academic researcher to the data provider. I will also share my personal experiences of three years of data requesting from multiple sources (academia, government and industry) and use of [clinicalstudydatarequest.com](http://clinicalstudydatarequest.com) and the SAS data access system.

### 4. Trish Groves (BMJ): "The future of data sharing"

The US Institute of Medicine (IOM) Committee on Strategies for Responsible Sharing of Clinical Trial Data published its discussion framework in January 2014 for public consultation. Many organisations responded, including the International Committee of Medical Journal Editors (ICMJE), whose interim guidelines laid the ground for journal policies.

Now that the IOM has concluded that data should be shared, and that it's time to move from the why to the what and how, how are ICMJE and journals planning to respond?

How will journals handle the big leaps in clinical trial transparency that are coming soon in the European Union and other territories?

And how are journals responding to industry's pioneering work on data sharing through initiatives such as the Yale Open Data





# PSI Conference 2015: Collated Abstracts

## Wednesday:

### “Prior” examples of Bayesian Analysis

This session takes a practical approach to Bayesian statistics. The three speakers will all be going through examples of where they have applied Bayesian methodologies to drug development. If you want to get some real hints and tips of running your own Bayesian analysis or are just curious then this is the session for you.

#### Speakers:

**1. Mark Belger (Eli Lilly): “The Development of a risk score through use of a Bayesian hierarchical model”**

Authors: Mark Belger, Karen Price, Fanni Natanegara – Eli Lilly  
In many applications of Bayesian models the use of prior information may be seen as a limitation, and so these models look to use non-informative priors. Through the example of developing a Risk Score for Stent Thrombosis, we demonstrate the benefits of incorporating informative priors.

Development of risk score models from individual trial or registry data often produce conflicting results as to the important factors, or are not large enough to identify all of the important factors needed to develop a risk score. The quality of these studies also varies considerably. The aim of the first part of this project was to synthesize the evidence from these studies and achieve a consensus view on the important risk factors and their associated weightings. In the second part, a systematic robust questionnaire was developed to capture the opinions of clinical experts and translate them to information that can be included in the Bayesian model, as informative priors. Using a Bayesian framework we were able to combine the data, from 44 studies, with the information from clinical experts, and achieve a consensus on the factors and their associated weightings to include in a tool to predict the risk of ST.

**2. Daniel Sabanes Bove (Roche): “Bayesian Learning in Early Phase Oncology: A Case Study”**

The early clinical stage of drug development is a learning phase: we are learning continuously about the drug's safety, pharmacokinetics, pharmacodynamics and efficacy, building on our current knowledge. Therefore Bayesian inference, with its

coherent concept of updating prior information with observed data to obtain the posterior information about quantities of interest, is a perfect match to early phase study designs and to broader clinical development questions.

This case study on a new biologic from Oncology starts with the entry-into-human phase I dose escalation study. It is shown how the modified Continual Reassessment Method (CRM) design incorporated reasonable prior assumptions about the expected safety profile, and ensured maximum flexibility for study conduct. A separate dose escalation was then planned for the combination with another new drug, with the design building on the two compound's information. As during the phase I it became apparent that a large proportion of patients developed anti-drug antibodies against the biologic, a small proof-of-concept study with a pretreatment aiming to diminish the immune response against the biologic was designed. Finally, the information gathered so far can be used to setup the entry-into-human phase I study for another biologic from the same platform.

The clinical development questions and Bayesian answers to them will be presented, with a focus on the decision making and practical considerations.

**3. Haijun Ma (Amgen): “Bayesian Hierarchical Modeling for Detecting Safety Signals in Clinical Trials”**

Detection of safety signals from routinely collected adverse event data in clinical trials is critical in drug development, but carries a challenging statistical multiplicity problem. Without multiplicity considerations, there is a potential for an excess of false positive signals. On the other hand, traditional ways of adjusting for multiplicity may fail to flag important signals too often. Bayesian hierarchical modeling is appealing for its ability to explicitly model AEs with the existing coding structure so that they can borrow strength from each other depending on the actual data, as well as moderate extreme findings most likely due merely to chance. We implement such a model for subject incidence (Berry and Berry, 2004) using a binomial likelihood, and extend it to subject-year adjusted incidence rate estimation under a Poisson likelihood. We compare the performance of the Bayesian models with other commonly used statistical methods for analyzing AE data. In addition, we offer some practical considerations in applying this Bayesian signal detection method.

### Biosimilars: Same, same, not different?

Whilst generic drug development is clearly established, biosimilar drug development is on the other spectrum. As biologic drugs become more common with highly complex manufacturing processes, the definition for “similar” is more complicated. This session aims to look at both the regulator and drug developers' perspectives when designing and implementing biosimilar clinical development plans, and the data needed to support filing this class of drugs.

#### Speakers:

**1. Frank Fleischer (Boehringer-Ingelheim): “Clinical development of a biosimilar – statistical issues and solutions”**

The pharmaceutical business is currently experiencing the fact that many new biological entities (NBE) in particular monoclonal antibodies are reaching the end of their patent duration. Prominent examples are infliximab, trastuzumab, rituximab or etanercept, mostly owning registrations in immunology and oncology. As these monoclonal antibodies are complex molecules, established regulatory guidance and pathways used for the development and registration of NCE generics do not apply. Therefore many discussions and unclarities on how to develop biosimilars to these NBEs are ongoing. This presentation aims at elucidating the open critical statistical-methodological aspects of biosimilar development. Real-life case studies will be presented covering

- Planning an efficient phase I bioequivalence design by using a group-sequential method.
- Adaptive and alternative approaches for the clinical equivalence trial
- Correctly evaluating risk differences in response rates for an equivalence setting adjusting for covariates
- Interaction on statistical methodological topics with different stakeholders from regulatory and academia

Thereby it will be illustrated how to plan for an efficient clinical development of a biosimilar and how to capitalize on innovative statistical methods in such a context.

### 2. Dominik Heinzmann (Roche): “Development challenges for biosimilars: A Statistician’s view”

Biosimilar development programs differ from traditional development programs in that the idea of the program is not to demonstrate efficacy and safety of the biosimilar product per se, but rather establish similarity to an existing and comprehensively characterized reference product. According to regulatory biosimilars guidelines, clinical similarity is to be established based on equivalence type of trials. They should be conducted in the most sensitive population such that differences - if they exist - can most easily be detected. Efficacy endpoints used to assess similarity need to be sensitive and may thus differ from traditionally used clinical endpoints. Once products are shown to be highly similar in the most sensitive setting, extrapolation into all approved indications of the reference product may be considered if scientifically justified. Various clinical and statistical challenges in assessing similarity will be discussed, including approaches to identify (most) sensitive population and endpoints.

### 3. Peter Volkers (Paul Ehrlich Institut - PEI): “Biosimilarity issues from a regulator’s perspective”

Biologicals, especially monoclonal antibodies, have improved the treatment of serious diseases in many areas. Over the next years patent protection for several of these products will expire. Biologicals are much more complicated than small-molecule drugs, thus the development of a biosimilar is more laborious than the development of a generic for a small molecule. In Europe the development of biosimilars is supported by a comprehensive regulatory framework dealing with quality, non-clinical and clinical aspects. While statistical aspects are often not explicitly mentioned in these guidance documents, statistics plays a vital role in assessing biosimilarity. Following an overview on the regulatory framework for biosimilars in Europe the presentation will briefly mention statistical issues related to analytical similarity. The main focus of the presentation will be on statistical issues to assess clinical similarity, discussing issues related to study design, endpoint selection etc.

## Contributed Papers Session: Challenges in Early Clinical Development

The following speakers have been selected from the contributed abstracts received to talk on topics related to challenges in early clinical development.

### Speakers:

#### 1. Trevor Smart (Eli Lilly): “Small PoCs - Are we expecting too much?”

There is always a drive for smaller and quicker studies, but how big should a Proof of Concept (PoC) study be? Over the years decision rules to pass PoC studies have changed from the traditional 2 sided 5% tests, to 1 sided 5, 10 or 20% tests or more recently Bayesian probability boundaries. This has in some cases resulted in smaller PoC studies. What should we expect from these small studies?

Two recent small neuroscience PoC studies are used as case studies to discuss whether the drive to go small has gone too far or whether their purpose should be viewed differently. Both studies yielded initial inconclusive results with potential issues. In the first study the results were on the pre-specified decision boundary. In the second there was imbalance in an important covariate between the treatment groups. For both studies exploratory analysis and biomarker data helped put the results in context to enable decisions to be made.

Were the studies too small, were we expecting too much or despite the issues, can we view them as successful in terms of making a PoC decision? When are small PoC studies appropriate? This will depend on where in the clinical plan risk was being removed. As single studies establishing the efficacy of the compound, the studies may be too small, but given the wider clinical plan and portfolio considerations these small studies can still have value in specific situations.

#### 2. Charles Warne (Roche): “The combination of randomized and historical controls in clinical trials: Methods to borrow dynamically from historical data based on compatibility between randomized and historical controls”

Authors: Charles Warne, David Dejardin, Paul Delmar, Katie Patel  
The incorporation of historical control data into the analysis of a clinical trial can be justified when it is difficult to recruit or unethical to randomise patients to a control arm, and the historical control patients can be considered equivalent to the clinical trial population. Bayesian methods offer a natural framework in which to dynamically borrow historical data based on their compatibility with the current data, and various approaches have been proposed in the literature. However, it remains unclear which of these methods is optimal when the estimate for the historical control is only based on a single study.

The primary concern linked to the inclusion of historical data in the analysis of a clinical trial is the impact on type I error and power in case the historical data are very different from the randomized control data. We report the results of a simulation study applied to a rare disease non-inferiority Phase III trial with a binary endpoint evaluating 1) different methods of dynamically borrowing from the data with respect to their ability to adjust for the compatibility and the impact on power and Type I error; and 2) design options to recover the lost power in case of incompatible historical data.

#### 3. Judith Anzuers-Cabrera (Roche): “Statistical considerations for modifying the design of a study that is already recruiting patients”

A phase II study within the asthma program started recruiting patients into two treatment arms. Three months after recruitment had started the competitive landscape and payer assumptions of the program changed, so the design of the study was revisited to account for these changes. The study team was faced with the challenge of adding a new dose arm to the study while recruitment was ongoing. This major change to the design presented different operational and statistical challenges that the team had to sort out. Among the statistical issues were (i) avoiding unblinding of sites already recruiting patients in the study, (ii) avoiding the introduction of bias and confounding to the study, (iii) changing the randomization algorithm taking into account the number of patients already recruited, and (iii) trying to maintain balance across the three treatment arms within the registry recruiting patients since the beginning of the study and the new sites that came on board after the new arm was added. In this presentation I will discuss how the team solved all the challenges, and present the results of a simulation study that allowed us to make decisions on how to change the dynamic hierarchical randomization allocation (DHRA) algorithm.

## Plenary Session: Joint HTA and Regulatory Advice

### Title: Engaging with Regulatory Agencies and Health Technology Assessment (HTA) Bodies in Scientific Advice

The majority of statisticians working in the Pharmaceutical Industry are familiar with the Committee for Medicinal Products for Human Use (CHMP) scientific advice process where sponsors of new medicines in development seek feedback from the CHMP





Scientific Advisory Working Party (SAWP) on their proposed clinical development programmes and clinical studies that will form the basis of a future Marketing Authorisation Application (MAA). Until recently, there was no mechanism to obtain feedback from several HTA bodies (such as NICE in the UK, HAS in France and IQWiG in Germany) in a co-ordinated manner on the clinical evidence being generated and enable HTA bodies to discuss key elements of clinical trial designs important for reimbursement submissions. A partnership has now been established between the European Medicines Agency (EMA) and HTA bodies in Europe that enable sponsors to seek scientific advice from both sets of stakeholders, either together in a joint forum, or separately in parallel meetings. The Shaping European Early Dialogue (SEED) Initiative was launched in late 2013 to identify case studies to pilot this new process. Draft guidance was issued in 2014 describing how this parallel scientific advice process could work, including identifying the responsibilities of sponsors, the EMA and HTA bodies. Experiences from the case studies will help to further refine the process.

The objectives for engaging early with the EMA and HTA bodies in Europe for scientific advice will be described by regulatory and HTA body representatives. Industry experience on the early pilots conducted will be shared and key lessons learnt from all stakeholders discussed. Reflections on whether the SEED initiative has been successful from a regulatory and industry perspective will also be given. The EFSPi/PSI HTA Special Interest Group recommends that statisticians involved in designing clinical development programmes and clinical trials are involved and participate in this process.

Chair: Chrissie Fletcher (Executive Director Biostatistics, Amgen)

### Speakers:

1. David Wright (Deputy Manager of Statistics and Pharmacokinetics Unit, MHRA)
2. Francois Meyer (Advisor to the President, International Affairs, HAS)
3. Michael Happich (HTA Director, Eli Lilly)



@PSlupdate



PSI (Statisticians in the Pharmaceutical Industry) Ltd



[www.psiweb.org](http://www.psiweb.org)