BAYES 2019

AGENDA

including

ABSTRACTS

Day 1 – Tuesday, 21st May 2019

13:30 - 14:00 Registration & Welcome coffee

14:00 - 18:00 **Course: *How to practically implement adaptive design?***
Adrian Mander and James Wason, MRC Biostatistics Unit, University of Cambridge

Including coffee break (15:45 – 16:15)

Day 2 – Wednesday, 22nd May 2019

08:00 - 08:40 Registration continued

08:40 - 09:00 Welcome to the main conference

Session 1	**Session Theme: Decision making in digitalized world**

Chair: Emmanuel Lesaffre

09:00 - 09:45 **Invited talk: *Prof. Kerrie Mengersen, Queensland University of Technology, Australia***

Bayesian Learning for Decision Making in the Big Data Era

09:45 - 10:05 *Marmaduke Woodman, Daniel Lee, Eric Novik*

Translational Neuroscience: from dynamical systems to personalized medicine

Over the past decade, we have shown that biologically realistic brain network models composed of mathematical dynamical systems models coupled with subject-specific structural information has explanatory power beyond each of the components separately. This approach has been successfully applied to explain healthy brain function as well as in clinical translation to explain stroke and epilepsy.

We apply this approach to epilepsy, reconstructing connectivity matrices of human epileptic patients from Diffusion Tensor weighted Imaging (DTI) data. Stability analyses of propagating waves quantifies the degree of epileptogenicity and predict conditions, under which seizures propagate to non-epileptogenic brain regions, explaining the responses to intracerebral electric stimulation in epileptogenic and non-epileptogenic areas. The dynamical system model in implemented and estimated in Stan, where we apply both full Bayesian inference and variational inference. We describe some of the computational challenges in implementing and estimating this model in Stan.

These results provide subject-specific guidance in the presurgical evaluation of epileptogenicity based on electrographic signatures in intracerebral electroencephalograms and have been validated in small-scale clinical trials. The example of epilepsy nicely underwrites the predictive value of personalized large-scale brain network models.

10:05 - 10:25 *Pantelis Samartsidis, Shaun R. Seaman, Silvia Montagna, Andre Charlett, Matthew Hickman and Daniela De Angelis*

A Bayesian multivariate factor analysis model for evaluating an intervention using observational time-series data on multiple outcomes

Researchers are often interested in assessing the impact of an intervention on an outcome of interest in situations where i) the intervention is non-randomised; ii) few units are treated; iii) the intervention is binary; and iv) there are outcome measurements at multiple time points, both before and after the intervention. Several methods for causal inference in this setting have been developed, e.g. the factor analysis (FA) model. However, these methods are applied to a single outcome. When there are multiple outcomes possibly affected by the intervention, modelling these outcomes together may provide a more precise estimate of the causal effect of intervention on any one of them.

In this work, we extend the FA model to handle multiple outcomes. Our model allows for sharing of information between the different outcomes by assuming that some of the latent factors are common. Moreover, by modelling each latent factor as an AR(1) process, it can make use of temporal correlations in the data. Finally, by taking a Bayesian approach, we account for the uncertainty in the number of factors.

Using simulation studies, we show that the proposed model can improve the precision of the intervention effect estimates and thus can lead to substantial gains in power. We apply our method to estimate the impact of stricter alcohol licensing policies on alcohol-related harms.

10:25 - 10:45 *Yingbo Wang*, Heinz Schmidli*, Arunava Chakravartty**

Bayesian non-inferiority design borrowing strength from historical controls with a meta-analytic-predictive approach

*Novartis Pharma AG, Basel, Switzerland

Non-inferiority trials are designed to show that a new treatment is not substantially worse compared to an active control. In this talk we present a case-study where a lower dose of an approved drug is compared for non-inferiority with respect to its approved dose. The goal of such a design in this study is to explore alternate dose regimens with better safety profile without compromising the efficacy in comparison to what has been seen for the approved dose. In a standard frequentist setting, such a study would require sufficient number of patients in the investigational and active control arm to ensure adequate power and type 1 error with respect to the non-inferiority margin. We propose an alternative method to assess the treatment similarity using a Bayesian framework. With meta-analytic approach (MAP) implemented in RbesT R package, we show how the historical active control data can be used not only to define the non-inferiority margin but also integrated together with the study data after taking into account the heterogeneity between the studies .

Such an approach allows for a more efficient allocation of the sample size between the active control and investigational arms. Using simulations we show the improvements in the design efficiency with respect to power and the control of type 1 error under different scenarios of clinical interest. We will present results to compare the Bayesian design with a classical non-inferiority design, and conclude with examples that extend this approach to multi-stage and multi-arm scenarios.

10:45 - 11:15 Coffee Break & Sweeties

11:15 - 11:35 *Alice Gosselin (Sanofi)*

Comparison Study for Matching Methodologies for Historical Data Borrowing

With an increasing need for accelerated development of drugs in areas where it may be difficult or even unethical to enroll patients in a control group, alternative ideas may be explored such as the use of historical data for inferential comparisons. There are two options of utilizing historical data: to augment a concurrent small control arm or to support a single arm study with comparison to a completely synthesized control arm. Because of the association between the baseline variables and the treatment effect of the investigational drug (predictive and / or prognostic), a precise matching procedure is required when using historical data. Without any matching, the imbalance in these variables between historical and concurrent patients would lead to potential bias due to confounding effects. Matching procedures allow the comparison groups to be as similar as possible in terms of baseline observables, after the trial is completed. Such analyses would have to be pre-planned in the Statistical Analysis Plan. As an example, the propensity score matching relies on the probability of treatment assignment conditional on observed baseline characteristics, or so called the propensity score. In this framework, conditional on the propensity score, distribution of observed baseline covariates are considered as similar, this reduce the impact of confounders. An alternative of this approach was proposed, called entropy balancing. It is a data processing method based on the moments of the covariate distribution that obviates the dependence of a parametric model and the need for continual balance checking by using a reweighting scheme that directly incorporates covariate balance into the weight function that is applied to the sample units. By simulation studies, we evaluate and compare the performance of the propensity score method and the entropy balancing method in terms of accuracy (bias reduction) and robustness (model dependency) in the treatment effect estimation.

11:35 - 11:55 *Lennon H. [1], Pittavino M.[2], Chajes V.[3], Plummer M.[4] and Ferrari P. [1]*

A Bayesian hierarchical model to integrate dietary exposure and biomarker measurements for the risk of cancer

[1] Nutritional Methodology and Biostatistics Group, International Agency for Research on Cancer (IARC), Lyon, France; [2] Research Center for Statistics, Geneva School of Economics and Management, University of Geneva, Geneva, Switzerland; [3] Nutritional Epidemiology Group, International Agency for Research on Cancer (IARC), Lyon, France; [4] Department of Statistics, University of Warwick, Coventry, UK.

In nutritional epidemiology, self-reported assessments of dietary exposure are prone to measurement errors. Estimates of the association between dietary factors and risk of disease can be biased. It was suggested to complement self-reported dietary assessments with objective measurements (i.e. dietary biomarkers). Dietary and serum biomarker measurements of industrial trans fatty acids from a nested case-control study within the EPIC multicenter study were integrated in a Bayesian model to explore the measurement error structure of the data, and relate dietary exposures of trans fatty acids to the risk of subtype specific breast cancer.

A Bayesian hierarchical model was developed, which included 1) an exposure model, to define the distribution of unknown true exposure (T); 2) a measurement model, to relate observed assessments, in turn, dietary questionnaires (Q), 24-hour recalls (R) and biomarkers (M) to T measurements; 3) a disease model, to estimate exposures/cancer relationship. The marginal posterior distribution is obtained from the joint posterior distribution using MCMC sampling techniques in JAGS.

The study included 397 and 1679 case/control pairs for ER- and ER+ breast cancer (BC), respectively. After mutual adjustment, T estimates of elaidic acid and omega-3 were inversely related to ER+ but not to ER- BC risk. In a previous study on kidney and lung cancers, which included 554 and 882 case/control pairs respectively, T estimates of folate and vitamin B-6 intakes were inversely related to kidney and lung cancer risk, with measurement error corrected parameter estimates consistently lower than estimates obtained using, in turn, Q and M measurements. Bayesian models offer powerful solutions to handle complex data structures.

11:55 - 12:05 180 seconds challenge:

- *Petra L Graham [1], John L Moran [2]*

Evidence Synthesis: a comparison of Bayesian and frequentist approaches

[1] Macquarie University, NSW Australia
[2] University of Adelaide, SA Australia

Frequentist two-arm meta-analysis of binary outcome data is one of the most frequently undertaken investigations in the biomedical literature. Although such meta-analyses of randomised controlled trials (RCT) generate estimates which have face-value credence, the same cannot be said for all approaches to meta-analyses which attempt to combine RCT and observational studies (OS). A naïve approach simply combines all such studies in a single analysis without accounting for differences in study type such as that of a recently published meta-analysis containing only 5 studies; 2 RCTs and 3 OS [1]. A slightly less naïve approach presents sub-group analysis based on type together with an overall pooled estimate. Bayesian methods, on the other hand, can account for differences in study type using a variety of approaches. One recently published method takes a model averaging approach utilising numerical integration rather than MCMC [2]. Alternatively, in a method proposed almost 20 years ago, the heterogeneity between study types as well as the heterogeneity within study types is directly modelled using a three-level hierarchical approach [3]. In this analysis we compare and contrast results from all of these approaches especially with regard to estimation with low study number.

References

1. Munshi L, Walkey A, Goligher E, Pham T, Uleryk EM, Fan E: **Venovenous extracorporeal membrane oxygenation for acute respiratory distress syndrome: a systematic review and meta-analysis**. *The Lancet Respiratory medicine* 2019, **7**(2):163-172.
2. Rover C, Wandel S, Friede T: **Model averaging for robust extrapolation in evidence synthesis**. *Stat Med* 2019, **38**(4):674-694.
3. Sutton AJ, Abrams KR: **Bayesian methods in meta-analysis and evidence synthesis.** *Statistical Methods in Medical Research* 2001, **10**(4):277-303.

- *Noel Cressie*, Lili Zhuang***

Bayesian forecasting of infectious diseases with SIRS models

*National Institute for Applied Statistics Research Australia (NIASRA), University of Wollongong, Australia
**Department of Statistics, The Ohio State University, USA

The susceptible-infectious-recovered (SIR) model and its extension, the SIR-susceptible (SIRS) model, have been used extensively to study the dynamical evolution of an infectious disease in a large population. SIR(S) models assume observed counts are “mass-balanced,” which means that the total observed count equals the sum of the observed counts of the individual components of the model. However, this does not take into account errors in the observations. Another challenge in SIRS-type modeling is to capture the stochastic or random nature of an epidemic process. In this research, we develop a Bayesian hierarchical SIRS (HSIRS) model to address both of these critical aspects of infectious-disease modeling, where we concentrate on the temporal variability of the count data. Through simulation, we compare the forecasting ability of the HSIRS model to the classical SIRS model, which is the model that assumes the observed counts are mass-balanced and the dynamical evolution is deterministic.

12:05 – 13:30 Lunch time

13:30 – 13:40 Serious Game

Session 2 Session Theme: Bayesian modeling

Chair: Laurent Estève

13:40 – 14:25 **Invited talk:** ***Prof. Jean-Michel Marin, University of Montpellier***
Bayesian model choice

14:25 – 14:45 *Amy Shi and Fang Chen (SAS institute Inc)*
Introducing the BGLIMM Procedure for Bayesian Generalized Linear Mixed Models

This talk introduces the new BGLIMM procedure, a high-performance, sampling-based procedure that provides full Bayesian inference for generalized linear mixed models (GLMMs). PROC BGLIMM models data from the exponential family distributions that have correlations or nonconstant variability; uses syntax similar to that of the MIXED and GLIMMIX procedures (the CLASS, MODEL, RANDOM, REPEATED, and ESTIMATE statements); deploys optimal sampling algorithms that are parallelized for performance; handles multilevel nested and non-nested random-effects models; and fits models to multivariate data. PROC BGLIMM provides convenient access, with improved performance, to Bayesian analysis of complex mixed models that you could previously perform with the MCMC procedure. We present the important features of PROC BGLIMM and show how to use it for estimation, inference, and prediction.

14:45 – 15:05 *Sophie Ancelet[1] & Gaëtan Gruel [2] & Eric Grégoire [2] & Aurélie Vaurijoux [2] & Laurence Roy [3] & Joan Francesco Barquinero Estruch [4]*

A Bayesian Poisson mixture model for model selection and dose estimation in biological retrospective dosimetry.

1 Institut de Radioprotection et de Sûreté Nucléaire (IRSN),
PSE-SANTE/SESANE/LEPID, Fontenay-aux-Roses, France;
2 IRSN, PSE-SANTE/SERAMED/LRAcc , Fontenay-aux-Roses, France;
gaetan.gruel@irsn.fr, eric.gregoire@irsn.fr;
3 IRSN, PSE-SANTE/SESANE, Fontenay-aux-Roses, France;
4 Autònoma de Barcelona, Facultat de Biociències, 08193 Cerdanyola del Vallès,
Espagne;

Scoring of dicentric chromosome aberrations in peripheral blood lymphocytes is considered to be the "gold-standard" biological method to estimate the radiation dose received by individuals after proven radiation accidents. In this specific context, two main questions arise: 1) "Given the number of dicentrics observed in some blood lymphocytes of a given individual, what are the estimated absorbed dose and its associated uncertainty?" and 2) "Was the radiation exposure total or partial?" Dose estimation is highly relevant to optimize patient-centered care and predict the health consequences of radiation exposure.

Moreover, dose estimation from dicentric counts can be crucial to clarify unclear radiation exposure scenarios. In this context, one important question is: 3) "Given the number of dicentrics observed, was the individual really exposed to ionizing radiation?" Frequentist statistical approaches are commonly used to answer the above questions that are then formalized as hypotheses testing and inverse regression problems but no consensus has been reached so far on the best way to proceed. In this work, we propose an alternative approach based on the Bayesian inference of a Poisson mixture model. This approach allows providing, in a unique and coherent framework, some rich and simultaneous probabilistic answers to the above three questions. Particularly, our mixture model is used as a Bayesian model selection tool that is relevant to answer questions 2) and 3).

A specific adaptive Metropolis-Hastings algorithm was implemented to avoid potential convergence difficulties when estimating the mixture weights. Using simulation studies and cytogenetic data from real radiation accident victims, we discuss the advantages of the proposed Bayesian approach compared to the classical ones. A sensitivity analysis to the prior choice on the unknown dose and the mixture weights was also performed.

Keywords. Bayesian statistics, mixture model, model selection, inverse regression, cytogenetics, biological dosimetry, ionizing radiations.

15:05 – 15:25 *Jincheng Zhou, James S. Hodges, M. Fareed K. Suri and Haitao Chu, University of Minnesota Twin Cities*

A Bayesian Hierarchical Model Estimating CACE in Meta-analysis of Randomized Clinical Trials with Noncompliance

Noncompliance to assigned treatment is a common challenge in analysis and interpretation of randomized clinical trials. The complier average causal effect (CACE) approach provides a useful tool for addressing noncompliance, where CACE is defined as the average difference in potential outcomes for the response in the subpopulation of subjects who comply with their assigned treatments. In this article, we present a Bayesian hierarchical model to estimate the CACE in a meta-analysis of randomized clinical trials where compliance may be heterogeneous between studies. Between-study heterogeneity is taken into account with study-specific random effects. The results are illustrated by a re-analysis of a meta-analysis comparing the effect of epidural analgesia in labor versus no or other analgesia in labor on the outcome cesarean section, where noncompliance varied between studies. Finally, we present simulations evaluating the performance of the proposed approach and illustrate the importance of including appropriate random effects and the impact of over- and under-fitting.

15:25 – 15:55 Coffee Break

Session 3 Session Theme: Non-linear modeling

Chair: Astrid Jullion

15:55 – 16:40 **Invited talk: *Dr. François Mercier, Roche***

Bayesian non-linear disease progression models

16:40 – 17:00 *Antoine Pissoort[1], Maud Hennion[1], Anne Bousseau[2], Caroline Denot[2], Bruno Boulanger[1]*

A Bayesian Disease progression model of Parkinson Disease combining RWD and natural history data to evaluate a new treatement

[1] Pharmalex, Mont-Saint-Guibert, Belgium
[2] Brainever, Paris, France

Parkinson disease (PD) is one type of progression (degenerative) disease that affects around 6 million people for which there is virtually no treatement existing. There is currently many projects around the world that attempt to develop new cures for this devastating pathology. One therapeutic approach in PD is to develop a treatement that either stop or slow down the disease progression. For that purpose the very aim of this statistical analysis is to demonstrate that the evolution of the disease for each treated patient is drifting from the expected natural trajectory given the past history of the patient and given the real world data available for PD patients.

The objective of this presentation is first to show how a Bayesian hierarchical piecewise non-linear PD disease progression model has been developed based on RWD available from the Parkinson's Progression Markers Initiative (PPMI) database. This model and those PPMI data have been used to elicitate priors for the same model to be used in a clinical trial and therefore used to optimize and size properly the clinical trial.

The clinical trial has first a observational natural history part (NHS) that last for one year on patients that will then enter into the interventional part that will also last one additional year. The Bayesian hierarchical disease model will be used on NHS and interventional data using (degenerated) priors from the PPMI data to evaluate changes in patient progression given their respective history.

The operating characteristics of the clinical trial and the related Bayesian model have been evaluated under different scenarios by means of extensive simulations. Specific inference strategies have been proposed using function of the model parameters to provide the scientists with meaningful and interpretable results about the potential efficacy of this new treatement in a proof-of-concept study with a rather limited number of patients.

The idea presented in this work can be extended to other progressive diseases, such as the Charcot disease (Amyotrophic lateral sclerosis). A small comparative analysis will be presented.

17:00 – 17:20 *Maud Hennion (1), Muriel Boulton (2), Bie Verbist (2), Bruno Boulanger (1)*

Pharmacokinetic Model with SAS - Proc MCMC with applications to preclinical pharmacology

(1) Pharmalex Belgium, Mont-Saint-Guibert, Belgium
(2) Janssen, Beerse, Belgium

Population PK/PD models describe the drug concentration-time course in body fluids resulting from administration of a certain drug dose. Compartment models, i.e. systems of differential equations, are typically used to conceptualize the mechanisms that take place in the interaction between an organism and a drug product. Software such as NONMEM is considered as the gold standard within the pharmacometrician community for population PK/PD modeling.

In recent years, the use of Bayesian methods has spread widely in many application sectors. While Bayesian applications in PK/PD modeling stay limited, statistical software such as SAS or STAN, are developing new functions for population pharmacokinetic modelling. The presentation will focus on the use of SAS – Proc MCMC and the recent features proposed and tips to be used.

In addition to the ODE solver already available in SAS, a new function (CMPTMODEL) is now proposed for Bayesian analyses to fit compartment models. Whereas the ODE solver allows to fit a system of differential equations in any context, the CMPTMODEL function is dedicated to pharmacokinetic applications.

The purpose of this talk is to present the different functionalities offered by SAS to fit differential equations in a Bayesian context. Emphasis will be placed on the pharmacokinetic modelling and based on examples, options and capacities of the new CMPTMODEL function will be presented.

As illustrative example the results from a preclinical study with sparse designs for supporting a PK/PD analysis will be shown and programming details and output explained.

17:20 – 17:40 *Fabiola La Gamba[1,2], Tom Jacobs [1,] Jan Serroyen [1,] Helena Geys [1,2], Christel Faes [2]:*
Bayesian sequential integration within a preclinical PK/PD modeling framework: Lessons learned

[1] Janssen Pharmaceutica NV, Turnhoutseweg 30, B-2340 Beerse, Belgium
[2] Center for Statistics, Interuniversity Institute for Biostatistics and statistical Bioinformatics, Agoralaan building D, B-3590 Diepenbeek, Belgium

Although Bayesian methods are expanding considerably in various scientific areas, their applications in the field of pharmacokinetic/pharmacodynamic (PK/PD) modelling and simulation is still relatively limited. In this work, Bayesian techniques are used to facilitate the estimation of a novel PK/PD model which is developed to quantify the extent of PD synergy between two compounds using historical in vivo data.

Since the data consist of a series of 11 trials performed sequentially, a Bayesian sequential integration is considered: the posteriors resulting from the analysis of one trial are used to specify the hyperparameters of the priors of the next trial. The recursive update of posterior distributions whenever new information is available is less computationally intensive compared to the analysis of all data up to the current trial. However, this method implies the analysis of a limited amount of information during the first integration steps, which may hinder the estimation process. The aim of the present work is to discuss challenges as well as opportunities which are related to the impact of (i) prior specification, (ii) random effect choice and (iii) experimental design. In addition, the results from an extensive simulation study assessing the performance of the Bayesian sequential integration for an increasing model complexity are evaluated.

The results suggest that the use of informative prior distributions reduces the correlation among parameters and improves the accuracy of estimates. Moreover, choosing the random effect on a parameter that is not highly correlated with others avoids overcompensations, thus ensuring better predictions. On top of that, trials should be designed so that each of them explores an exhaustive number of doses and sampling times.

17:40 – 17:50 Day 2 closure & Social event details

19:30 – 23:00 Social Event

Day 3 – Thursday, 23rd May 2019

08:00 - 08:20 **Registration continued**

08:20 - 08:30 Welcome to Day 3

Session 4 Session Theme: Designing clinical trials

Chair: Heinz Schmidli

08:30 - 09:15 **Invited talk: *Dr. Kaspar Rufibach, Roche***
Bayesian Predictive Power: Theory, challenges in implementations and perspectives

09:15 – 09:35 *Arnaud Monseur [1] , Karine Imberdis [2] , Didier Poirault [3]*

From Power and Assurance to Bayesian Power: application to probability of success

[1] *Pharmalex Belgium-Mont Saint-Guibert Belgium*
[2] *Sanofi- Montpellier, France*
[3] *BioMérieux-Marcy L'Etoile, France*

In clinical early phase Development, it has been asked frequently to estimate the probability of success (PoS) of a second study taken into account the data of a first study as for example with a BA study (CO study with N around 20, to test several formulations) followed by BE Study (CO study sized to prove the Bioequivalence with the formulation selected vs the reference formulation, N around 100) or a FIH study followed by a POC study. We propose here several POS estimations obtained from power and assurance to Bayesian power on a real case of a future similarity study with 188 patients based on a FIH data study observed with 24 patients.

Power, the probability to correctly reject the null hypothesis, is a primary statistical concept which lies at the heart of statistical analyses even when Bayesian estimations are used. One of the main caveats of power computation is that it is a conditional probability and is not acknowledged as such. Assurance however, as defined by O'Hagan et al. [2005], enables the statistician to set themselves free of this conditionality of a pre-defined effect size. Indeed, assurance is the expectation of power over all possible effect sizes. More recently, bayesian power, as for example defined by Kruschke [2014] has also gained in importance and consists in an equally important measure.

This talk proposes to expose the differences and similarities between these measures and to implement the measures on the real case considered. The aim is to familiarize the statistician to alternative methodologies that may allow to produce more reproductible and less biased results. Said otherwise, relieving the practicing statistician from the mysticism of the frequentist paradigm and welcoming measures that do not rely on implicit conditionality.

09:35 – 09:55 *Astrid Jullion, Lisa Hampson*

Probability of success: a viable concept to inform the early development of a topical drug?

At key milestones in the development of a new medicine, it can be useful to calculate risk metrics measuring the chance that future trials will achieve their success criteria. At Novartis, we are developing a more quantitative approach for calculating the probability of success (PoS) of a program at the end of phase II which leverages internal clinical data, cross-industry historical success rates, and expert opinion or external data if needed. Using a Bayesian approach seems natural since it allows one to account for uncertainty about the size of treatment effects and the relevance of the phase II data when there are differences between development stages.

It is therefore appealing to apply the same concept in early drug development. However, additional challenges arise such as smaller sample sizes in phase I and potential changes in population, and/or formulation between phase I and phase 2,... Propogating these additional uncertainties through the PoS calculation is key to truly represent the risk associated to the decision.

In this work, we will assess whether the PoS of the phase IIa/b calculated after a phase I study of a topical drug in atopic dermatitis adult patients will be informative enough to support the further development of the compound. The additional sources of uncertainty include:

- the vehicle effect that can be higher in a well-controlled phase I environment - the age of the population, the target population being children - the formulation as technical team may come with an improved one later in the development.

The objective of the presentation is to share the journey to set-up the concept of PoS in early development and to summarize what reasonable expectations the clinical teams can have with regards to its utility to support their decision making.

09:55 – 10:15 *Samuel Pawel, Leonhard Held*

Predictive Evaluation of Replication Studies

The conduct of replication studies not only plays an important role in assessing the credibility of scientific discoveries, but is also often a regulatory requirement in the case of clinical research. Throughout the last decade, the so-called reproducibility crisis has stimulated many researchers from different fields to conduct large-scale replication projects. By now, some of these projects have been completed and their data made available to the public.

Using data from four different replication projects, we computed probabilistic forecasts of the replication studies' outcomes based on frequentist predictive, prior predictive, and shrinkage predictive distributions. The amount of shrinkage was estimated using empirical Bayes, depending on the amount of evidence in the original study. We then assessed these predictions with respect to discrimination, calibration and sharpness using established methodology such as proper scoring rules, probability integral transform, calibration slope, and area under the curve.

Although there were differences between the different fields, the forecasts based on shrinkage predictive distributions generally showed the best predictive performance. Our results suggest that many of the estimates from the original study were too optimistic, which may have been caused by publication bias or questionable research practices. Moreover, our results also indicate that the use of conventional statistical significance as criterion for replication success may be questionable since from a predictive viewpoint, non-significant replication outcomes are often still in agreement with a significant result from the original study.

10:15 - 10:45 Coffee Break & Sweeties

10:45 – 11:05 *Charlotte Micheloud, Manuela Ott & Leonhard Held*

Sample Size Calculation for Replication Studies

University of Zurich, Epidemiology, Biostatistics and Prevention Institute, Hirschengraben

As a consequence of the so-called `replication crisis' [Ioannidis, 2005], an increasing number of replication studies have been conducted to determine the reliability of the original _ndings. Ideally the procedures of the replication study should be as closely matched to the original study as possible. How-ever, selecting the same sample size in the replication study as in the original study can lead to a severely under-powered design and true e_ects may not be detected. Furthermore, using standard sample size calculations is not well suited because the uncertainty of the original e_ect estimate is ignored.

One way of tackling this issue is to incorporate a normal prior centered around the original e_ect estimate and with variance inversely proportional to the original sample size [Spiegelhalter et al., 2004]. We adapt the use of Bayesian methods in power and sample size calculations for sequential trials [Dallow and Fina, 2011, Ru_bach et al., 2016] to the setting of replication studies. It turns out that Bayesian methods generally lead to larger sample sizes than standard methods to reach a certain level of power. Furthermore, the resulting `Bayesian power' tends to (1 minus the one sided p-value of the original study) as the replication sample size increases. Adding more subjects to the replication study can even decrease the predictive power if the p-value of the original study is only 'suggestive', i.e. only slightly below the signi_cance level. We illustrate these properties using data from the Open Science Col-laboration project on the replicability of psychological science [Open ScienceCollaboration, 2015].

References

Nigel Dallow and Paolo Fina. The perils with the misuse of predictive power.Pharmaceutical Statistics, 10(4):311{317, 2011. doi:10.1002/pst.467.

John PA Ioannidis. Why most published research _ndings are false. PLoS Medicine, 2(8):e124, 2005. doi:10.1371/journal.pmed.0020124.

Open Science Collaboration. Estimating the reproducibility of psychological science. Science, page aac4716, Sep 2015. doi:10.1126/science.aac4716.

Kaspar Ru_bach, Hans Ulrich Burger, and Markus Abt. Bayesian predictive power: choice of prior and some recommendations for its use as probability of success in drug development. Pharmaceutical statistics, 15(5):438{446, 2016. doi:10.1002/pst.1764.

David J Spiegelhalter, Keith R Abrams, and Jonathan P Myles. Bayesian Approaches to Clinical Trials and Health-Care Evaluation, volume 13. John Wiley & Sons, 2004.

11:05 – 11:25 *Annette Kopp-Schneider, Silvia Calderazzo, Manuel Wiesenfarth*

Use of external information in clinical trials: What can be gained in terms of frequentist power?

The research is motivated by a basket trial in precision medicine in which adults with a specific molecular tumor profile are treated with targeted therapy and response to therapy is assessed. The population of children with this specific molecular profile is too small to warrant a separate pediatric trial. This motivates the implementation of a pediatric stratum in the adult trial and the setting suggests that information from the adult trial should be used for the pediatric stratum as "historical information".

An overview of different methods for borrowing from historical data has been given by Viele et al. (2014). Several adaptive methods have been proposed that dynamically discount the amount of information borrowed from historical data based on the conformity between the historical and current data. Adaptive power priors represent one of the approaches suited for this situation where the discounting factor can be selected by, e.g., an empirical Bayes approach as suggested by Gravestock et al. (2017). Another approach is the use of robust mixture priors as proposed by, e.g., Schmidli et al. (2014).

We will show that even in case of dynamic borrowing no power can be gained when strict frequentist type I error control is required. We exemplify this finding in the case of the pediatric arm of an adult trial and a dichotomous outcome for various methods of dynamic borrowing. We discuss that this counter-intuitive limitation is true in any situation in which a uniformly most powerful test exists and show situations for which this applies.

However, if prior information is reliable and consistent with the data generating process, the final trial decision will be associated with a lower error. Thus, the key question should rather be if prior information is to be trusted, rather than if borrowing is beneficial for any possible true parameter value.

References:

Gravestock, I., Held, L.; COMBACTE-Net consortium (2017). Adaptive power priors with empirical Bayes for clinical trials. Pharm Stat. 16(5): 349-360.

Schmidli H., Gsteiger S., Roychoudhury S., O'Hagan A., Spiegelhalter D., Neuenschwander B. (2014) Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics 70(4):1023-1032.

Viele, K., Berry, S., Neuenschwander, B., Amzal, B., Chen, F., Enas, N., Hobbs, B., Ibrahim, J.G., Kinnersley, N., Lindborg, S., Micallef, S., Roychoudhury, S., Thompson, L. (2014) Use of historical control data for assessing treatment effects in clinical trials. Pharmaceutical Statistics 13(1):41--54.

11:25 – 12:10 **Invited talk:** ***Prof. Dr. Eric-Jan Wagenmakers, Amsterdam University, NL***
Bayesian benefits for the pragmatic biostatistician

12:10 – 13:25 Lunch time

13:25 – 13:35 Serious Game

Session 5 Session Theme: Inference in clinical research

Chair: Bruno Boulanger

13:35 – 13:55 **180 seconds challenge**

- *S. Calderazzo, M. Wiesenfarth, A. Kopp-Schneider*
  Bayesian clinical trials design and evaluation: a decision-theoretic view

Bayesian clinical trials allow taking advantage of relevant external information through the elicitation of prior distributions, which influence Bayesian posterior parameter estimates and test decisions. The impact of prior specification on frequentist operating characteristics is generally investigated at the design stage of the trial. More recently, concepts such as Bayesian power and type I error have been gaining increasing attention, allowing to average frequentist error rates according to hypothesized prior distributions under the null and alternative hypotheses (see e.g. Psioda and Ibrahim, 2018). Nevertheless, a global approach to evaluate the performance of different designs, which takes into account the relative costs of test error rates, estimation error, and of sampling, seems to be lacking. A Bayesian decision-theoretic viewpoint represents the natural framework to approach this task. In addition to the mean squared error and the sample size, Bayesian type I and type II error can be explicitly added in the integrated risk, and a test decision can thus be taken to minimize their sum (see Pericchi and Pereira, 2016). Fully Bayesian sensitivity analyses can also be performed by making a distinction between the prior of the data-generating process, and the analysis prior adopted to fit the data, a distinction adopted also in e.g. Psioda and Ibrahim (2018). We explore sensitivity of the integrated risk, the operating characteristics and of the optimal sample size for different analysis priors, including the increasingly popular empirical Bayes power prior (Gravestock and Held, 2017) and robust mixture prior (e.g. Schmidli et al., 2014).

Gravestock, I. & Held, L., Pharm. stat., 16(5):349–360 (2017).
Pericchi, L., & Pereira, C., Braz. J. Probab. Stat., 30(1):70–90 (2016).
Psioda, M. A., & Ibrahim, J. G., Biostatistics (2018).
Schmidli, H. et al., Biometrics, 70(4):1023–1032 (2014).

- *Maha Alsefri, Ruwanthi Kolamunnage-Dona, Maria Sudell*
  Bayesian joint modelling of longitudinal and time to event data: A review of methods and issues

In clinical research, there is an increased interest in joint modelling of longitudinal and time to event data since it is reduced the bias of parameter estimation and increase the efficiency of statistical inference. Inference and predication using this type of model are carried by using a frequent estimation or Bayesian estimation. The latter approach is overweight the former one since it is offered to fully capture the uncertainty of parameter estimate in the posterior distributions and does not need an asymptotic theory to drive the standard error. Also it can easily incorporate information from previous studies through parameter prior specification. A comprehensive review is undertaken on univariate and multivariate joint models. The review is focused on modelling assumption, association structure, Bayesian estimation algorithm, dynamic predication, software implementations, and clinical applications of the methodology.

- *Audrey Larue-Triolet (BioMérieux)*
  Bayesian method to predict the probability to comply to the criteria for a future study or to monitor a new study

The development process of a new diagnostic test is organized in several steps. Some performance studies of the diagnostic test (precision, detection capability, …) are performed several times during the different steps and must comply to predefined acceptance criteria to keep going. This is the case for the clinical sensitivity and specificity performances. The idea is to capitalize on preliminary studies to estimate the probabilities to comply to the acceptance criteria on these clinical performances, given the future protocol characteristics and using Bayesian framework. The goal is also to monitor the data during the enrollment of the new study.

Real data are used to demonstrate how Bayesian approach could simply be useful in the decision making during the development process of a diagnostic test.

13:55 – 14:40 **Invited talk:** ***Prof. Sara Geneletti, London School of Economics, UK***
Causal inference in clinical research

14:40 – 15:00 *Burak Kürsad Günhan [1,2], Tim Friede[1]*
Model-based meta-analysis using arm-based models

[1]Department of Medical Statistics, University Medical Center Goettingen, Goettingen, Germany
[2]Correspondence to: Burak Kürsad Günhan; email: burak.gunhan@med.uni-goettingen.de

Model-based meta-analysis (MBMA)[1] is a generalization of the "standard" meta-analysis in which dosing information from different studies is taken into account. MBMA has become an increasingly important aspect of drug development to inform study design and decision-making.[2] A Bayesian approach to fit generalized linear mixed models for the purpose of conducting MBMA is very attractive.

The aim of MBMA is estimating the dose-response relationship (e.g. Emax or logistic) by taking into account possible between-trial heterogeneity in the shapes of the dose-response curves. At present, model-based meta-analyses are commonly based on a *contrast-based model*,[3] which focuses on estimating relative treatment effects in comparison to a references dose such as placebo, and does not account appropriately for the variability in the estimates of the reference dose which is crucial for the estimation of the dose-response function. Inspired by the network meta-analysis literature,[3] we suggest the use of an *arm-based model* to fit MBMA models, in which absolute treatment effects are estimated.

We illustrate the proposed arm-based model, and compare it with the contrast-based model using a meta-analysis of a binary endpoint in migraine pain relief.[2] Simulation studies are conducted to assess the operating characteristics of the arm-based and contrast-based models under different scenarios.

Bayesian implementations of the arm-based and contrast-based MBMA models will be included in our publicly available **R** package MetaStan (https://CRAN.R-project.org/package=MetaStan).

Keywords: Model-based meta-analysis, dose-response, Emax, arm-based model, Stan

References

[1] Boucher M, et al. The many flavors of model-based meta-analysis: Part I–Introduction and landmark data. CPT: Pharmacometrics & Systems Pharmacology. 2016;5:54–64.
[2] Mawdsley D, et al. Model-based network meta-analysis: A framework for evidence synthesis of clinical trial data. CPT: Pharmacometrics & Systems Pharmacology. 2016;5:393-401.
[3] Dias S, et al. Absolute or relative effects? Arm-based synthesis of trial data. Research Synthesis Methods. 2016;7:23–28.

15:00 – 15:20 *Mark Wymer, Haitao Chu,Qinshu Lian, Jing Zhang*

Bayesian joint meta-regression methods adjusting for post-randomization variables in network meta-analysis ?

Though there is board consensus on accounting for study-level covariates in mixed treatment comparisons through network meta-regression, it is much more challenging to adjust for post-randomization variables, which are expected to differ between treatment arms within a study. Examples include differential noncompliance, measured as the proportion of premature treatment discontinuation or drop out, loss to follow-up, or change to an alternative therapy. In existing network meta-regression methods, study-level covariates are assumed to be fixed. However, post-randomization variables are generally considered random and thus cannot be adjusted for by existing methods. In this paper (talk), we will propose novel Bayesian joint network meta-regression methods to account for post-randomization variables, which enables more accurate estimation of treatment effects. We will illustrate the proposed methods through simulations and real data analyses.

15:20 – 15:40 *Kelly R. Moran, Elizabeth L. Turner, David Dunson, Amy H. Herring*

Bayesian hierarchical factor regression models to infer cause of death from verbal autopsy data

It is often of critical interest to determine and study the cause of death (COD) for individuals and the cause-specific mortality fraction (CSMF) for populations. Post-mortem autopsies, considered the gold standard for COD assignment, are generally not performed when individuals die outside the hospital. In many settings, even when individuals die in the hospital, full autopsies are difficult or impossible to implement due to expense or cultural norms. Thus data on COD is often scarce or unrepresentative of the entire population of deaths, and the CSMF is often poorly estimated. For this reason, Verbal Autopsies (VAs) are commonly conducted, consisting of a questionnaire recording demographic information, known medical conditions, symptoms, and other factors for the recently deceased. VA data consist of multivariate mixed scale observations for each individual, with COD directly measured for a subset of labeled individuals. The goal of this work is to improve modeling of VA data to infer more accurately COD and CSMF, while providing accurate uncertainty quantification in associated inferences. With this goal in mind, we propose a novel class of hierarchical factor regression models that avoid assuming conditional independence in the VA measurements given COD. Our approach allows both the mean and covariance to vary with COD category and can include covariate information on the decedent (such as age/gender), region (e.g. malaria endemicity or spatial location), or event (e.g., season, year). Taking a Bayesian approach to inference, we develop an MCMC algorithm and validate our FActor Regression for Verbal Autopsy (FARVA) method in simulation experiments. An application of FARVA to VA data in Tanzania shows improved goodness-of-fit and better predictive performance in inferring COD and CSMF over competing methods. We offer both customizable code for statisticians, and automatically produced visualizations for use in communication of results to physicians and governmental authorities.

15:40 - 16:10 Coffee Break & Sweeties

Session 6 Session Theme: Bayesian Statistics in CMC

Chair: Didier Poirault

16:10 – 16:55 Invited talk: *Dr. Mike Denham, GSK, UK:*

Bayesian Statistics in CMC

16:55 – 17:15 *Piotr Juszczak and Fernando Garcia-Alcalde*

On supporting development of diagnostic assays with Bayesian modelling and simulation

Advanced Analytics Group, Roche Diagnostics, Basel, Switzerland

OBJECTIVE To introduce how Bayesian modelling is used to support design and development of diagnostic assays.

BACKGROUND A team that develops a diagnostic assay usually comprises of biologists, chemists, engineers, and quantitative experts. The task of a quantitative expert, in the team, is to build a quantitative model, that models a physical or bio-chemical phenomenon that generates a measured signal. This signal is proportional to an amount of a substance that we would like quantify. The signal generating process is a non-linear and longitudinal in nature and it is typically modelled by a system of ODEs or PDEs. These models can be implemented either in Bayesian or Frequentist frameworks, however prior knowledge, i.e. a factory specification range and a variability of the assay components are known, therefore Bayesian framework, that incorporates this additional information, is more suitable here. As a mechanistic part of the data generating process is assumed to be known, one of the tasks of quantitative expert is to find sources of variability, that can influence measurements, e.g.: pH, temperature, hematocrit, other proteins, collection or preparation of a sample, to mention a few. In addition, based on the build Bayesian model the quantitative expert can support further improvement of the diagnostic device by selecting the device components, e.g. different antibodies or sensors, that will maximise the signal to noise ratio. All presented models have been implemented in Stan or Monolix.

CONCLUSIONS During development of the diagnostic device, a mechanistic part of a signal generating process is assumed to be known, and most of prior knowledge, i.e. a factory specification range and a variability of the assay components are known. Therefore, Bayesian modelling and simulation framework is a suitable tool, in hands of a quantitative expert, to support design and improvement of diagnostic devices.

17:15 – 17:35 *Bernard G Francq, Dan Lin, Mikaël Le Bouter, Mélissa Scozzari, Walter Hoyer (GSK Vaccines)*

Bayesian Coefficients of Variation in Linear Mixed Models, Random Effects and Precision in Assay Qualification

During development of a vaccine, different analytical methods for determining the antigen concentration or the (relative) potency in the produced vaccine batches, as well as for determining the titer of antibodies in blood samples, need to be developed. Accuracy of the assay is assessed to confirm the analytical method characteristics are aligned with the objective of its usage. Accuracy is the closeness between the reference value and an individual measured value, taking into account the systematic error (trueness) and random error (precision). In an Reproducibility and Repeatability study, a linear mixed model across all samples is applied to evaluate the precision of the assay through the estimate of variance components (repeatability and intermediate precision). Furthermore, the Coefficient of Variation (CV) is the quantity used to express the intermediate precision and repeatability of an assay through the whole range of values. CVs are the relative standard deviations, calculated per variance component and by or across concentrations within a mixed model. Calculating Confidence Intervals for CVs is challenging as the uncertainties of both the fixed effects and of (the sum of) variance components must be taken into account.

The advantages and convergence of Bayesian credible intervals for CVs will be discussed in the framework of linear mixed models with different sample sizes. The literature about CV in mixed models is scarce as often the methodology is developed for univariate design, which is not appropriate for unbalanced or more complex designs.

We compare the frequentist approach (using the McKay method) to the Bayesian approach under a wide variety of designs with a variance component structure (random, nested, crossed). Both statistical approaches will be evaluated by means of simulations and application to a case study in assay qualification.

17:35 – 17:55 *Xiaolong Luo, Jun Zhang, Tianlei Chen, Judy Li, Mike Branson (Celgene Corporation)*

Comparison of Bayesian Multivariate Models for Detecting Safety Signals in Clinical Trials

Identification of safety signals from clinical trial adverse event data is important and challenging. It plays a significant role in the benefit and risk assessment of an experimental product. It is complicated due to numerous adverse events that can occur and are potentially statistically related. Several Bayesian models have been proposed in the last decade to approach the problem, including hierarchical-normal and Ising models, to address the correlation among different adverse events and the intrinsic complexity of interpretation due to the multiplicity of results. We implemented these models for an integrated safety database from multiple clinical trials and proposed a standardized measure to compare their performance. We also evaluated the reliability using training and test datasets. In this talk, we will present the results, discuss some limitation of these methods, use forest plots to visualize the adverse event data, and share our experience in fine tuning and selecting methods to characterize safety signals and help inform our understanding of clinical trials safety data.

17:55 **BAYES2020 Announcement**

Day 4 – Friday, 24th May 2019

08:00 - 08:20 Registration continued

08:20 - 08:30 Welcome to Day 4

Session 7 Session Theme: Bayesian statistics for 21st century cures

Chair: Leonhard Held

08:30 – 09:10 ***J. Jack Lee, Ying Yuan, Department of Biostatistics, University of Texas MD Anderson Cancer Center***

Bayesian Model-Assisted Designs and Their Applications for Early Phase Clinical Trials – When Simplicity Meets Superiority

Many novel adaptive designs have been published in the last two decades. Although these novel adaptive designs possess good statistical properties, most of them have not been widely implemented in real trials. Three major impediments are (1) complicated statistical modelling, (2) demanding computations, and (3) hard-to-be-understood by clinical researchers. We introduce and review a new class of novel adaptive designs, known as model-assisted designs, to remove these hurdles and to facilitate the increasing use of novel designs to improve the efficiency and success of clinical trials. Model-assisted designs combine the transparency and simplicity of the conventional algorithm-based designs with the superiority and rigorousness of model-based designs. They enjoy the superior performance comparable to more complicated, model-based designs, but can be implemented as simple as the conventional designs. A few model-assisted designs will be discussed including the Bayesian optimal interval (BOIN) design, the keyboard design, the time-to-event BOIN (TITE-BOIN), the BOIN combination design, and the Bayesian Optimal Phase 2 (BOP2) design, etc. Similar to all trial designs, the design parameters need to be carefully chosen and calibrated through simulation studies to ensure desirable operating characteristics. Freely available Shiny applications are provided to facilitate the adoption of model-assisted designs. Model-assisted designs establish a new KISS principle: Keep It Simple and Smart!

09:10 – 09:30 *Forrest Williamson (Eli Lilly and Company)*

Bayesian applications in rare disease clinical trial research

Rare diseases have many legal definitions worldwide, but ultimately it is a disease with relatively low prevalence. The traditional archetype of multi-stage drug development, the end of which often consists of multiple fully powered, randomized, and controlled clinical trials, is often infeasible for many rare diseases -- by definition there are not many patients to study. It is essential to leverage all available information to understand the disease state (natural histories, real world evidence, registries) and potential impact of therapeutic intervention (external controls, borrowing on effect in other populations). Bayesian statistics provides a natural structure for incorporating this information in to the analysis of a new trial. This talk will introduce some of the Bayesian approaches that may be considered in rare disease states, and provides examples in cognitive degeneration and paediatric extrapolation.

09:30 – 09:50 *Barbara Wendelberger[1], Melanie Quintana[1], Julie Sapp[2], Leslie Biesecker[2], Scott Berry[1]*

Guiding clinical trial design for a rare disease using natural history data and Bayesian disease progression modeling

[1]Berry Consultants, Austin, TX, USA
[2]National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA

Rare diseases often present unique challenges to clinical trial design, which may include small, heterogenous patient populations, insufficient understanding of disease etiology, and poorly developed study endpoints. Proteus syndrome is a rare condition resulting from a mosaic genetic mutation that manifests in the overgrowth of skin, bone, and other tissues. Plantar cerebriform connective tissue nevus (CCTN) is its most specific symptom. Here, we investigate the design of a single arm clinical trial in Proteus syndrome through 1)exploration of natural history data, 2)disease progression modeling of CCTN, 3)construction of a virtual patient simulator, 4)integration of a treatment effect, and 5)analysis of CCTN as a primary endpoint using a Bayesian disease progression model (DPM). This approach allows us to understand disease progression, simulate virtual patients in virtual trials, and investigate our primary analysis model across a range of treatment effects. Importantly, the Bayesian DPM provides flexibility that allows for differential patient follow up in terms of both number and timing of patient visits. This flexibility enables the DPM to estimate the control rate of disease progression from the CCTN natural history data and also estimate the rate of disease progression relative to the natural history rate of progression using CCTN data from treated subjects in the single arm trial. Designing an effective single arm trial that leverages information from a previous natural history study requires dedicated collaboration between clinicians and statisticians. The Bayesian DPM encapsulates both clinical disease progression and statistical disease modeling and provides an intuitive framework for the explanation and interpretation of results for regulatory authorities, effectively providing solutions for researching a challenging rare disease.

09:50 – 10:10 *Angély Loubert[1], Marie-Laure Delignette-Muller[2], Fabien Subtil[1], Muriel Rabilloud[1]*

Combination of prior distributions elicited from expert opinions previous to Bayesian inference. Application to personalized medicine.

[1] Service de Biostatistique - Bioinformatique, Pôle Santé Publique, Hospices Civils de Lyon, F-69003, Lyon, France ; Université de Lyon, F-69000, Lyon, France ; Université Lyon 1, F-69100, Villeurbanne, France ; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Évolutive, Équipe Biostatistique-Santé, F-69100, Villeurbanne France

[2] Université de Lyon, Université Lyon 1, CNRS, VetAgro Sup, UMR 5558, Laboratoire de Biométrie et Biologie Evolutive, 69622 Villeurbanne, France

In clinical trials in the context of personalized medicine, it may be difficult to enroll enough subjects to conclude on the tested treatment efficiency. One option to increase the power of the study is to bring additional information through elicitation of expert opinions about the primary endpoint value for each arm of the trial. This information expressed as a prior distribution is combined with the information bring by the trial results in a Bayesian inference approach. The challenge is to find a way of combining several distributions from different experts to build a valid prior distribution. The work compares two ways: i) Mathematical combination methods using either linear or logarithmic pooling; ii) Modelling with a supra-Bayesian approach. The latter approach assumes that the distribution of each expert stems from a single general distribution and allows to take into account the uncertainty of each expert, as well as a potential variability between experts. The objective of this work is to compare the different approaches, in order to better understand their impact on the prior information delivered in the context of clinical trials. This work uses a combination of eight distributions elicited from 8 physicians regarding the proportion of patients in non-progression after a six-week treatment (A vs. B) for an Ear-Nose-Throat epidermoid carcinoma. To better characterize the properties of the different approaches, they are also applied on simulated expert distributions, following various scenarios where number of experts and variability among their opinions vary. Results from both simulated and real elicited data combination studies are discussed, and interpretation of the different approaches is provided.

10:10 - 10:40 Coffee Break & Sweeties

10:40 – 11:00 *Pierre Colin (Sanofi)*

Use of Probability of Success - A way to manage more available information

The Probability of Success (PoS) is more and more widely used in drug development. The main advantage compared to the statistical power is to take into account the uncertainty about the true treatment effect (or any parameter of interest). The PoS can be provided to support decision-making (Go/noGo decision, futility/overwhelming stopping rules, …) in an ad hoc process or in a predefined decision rule. In this article, we focus on gathering information from a prior study and a blinded interim analysis, during a randomized two-arms clinical study, to provide the PoS of the final analysis based on a binary endpoint. The blinded information (a pooled estimator of event rate and a failed overwhelming efficacy stopping rule) provides a strong information on the upper bound of the treatment effect. An updated estimate of treatment effect is obtained through the Bayesian framework, using the prior distribution (from a previous study) and the likelihood of the blinded information. This leads to an interesting PoS of the final analysis that could be used to support the decision to continue or stop the clinical study. Generally, this example highlights the difficulty of stopping rules definition and its impact on design operating characteristics.

11:00 – 11:20 *Anna Pöhlmann[1,2] Oliver Sailer[2]*

An R Shiny App to design and analyse basket trials in Oncology

[1] *University of Bamberg, Bamberg, Germany*
[2] *Boehringer Ingelheim Pharma GmbH & Co. KG*

Keywords: Basket design, Bayesian Hierarchical Model, Go/No-Go decision, Oncology

In early clinical development in oncology, researchers often investigate a single treatment for multiple indications. To make decisions about further development for specific patient cohorts, basket trials are used to evaluate various subpopulations in a single trial for early signs of efficacy.

We introduce an interactive R Shiny App (Chang et al., 2018) that supports Statisticians in their interaction with Physicians in planning basket designs for response endpoints. The final design results in sufficiently large probabilities of correct Go and No-Go decisions per cohort and on program level.

The Shiny App simulates the whole process from recruitment to final analysis. The recruitment process is dynamic since start date, recruitment rate and number of patients can vary across cohorts. If data meet predetermined conditions, expansions of cohorts are possible. Futility analyses determine early No-Go decisions on cohort level based on predefined rules. Futility stops result in recruitment stops that influence the recruitment and final analyses of other cohorts. Every time a cohort is completed, all evaluable patients in all cohorts are included in the analysis model. A modification of the Bayesian Hierarchical Model (BHM) proposed by Berry et al. (2013) is used for analysis to allow for borrowing of information across cohorts.

We present an application of the Shiny App to a recent study that evaluates seven different subpopulations of patients with futility analyses and expansion cohorts.

Implementing Go/No-Go decisions for basket designs in a Shiny App helps to plan the study more efficiently. Its interactive nature facilitates discussions between Statisticians and Physicians.

References

Berry SM, Broglio KR, Groshen S, and Berry DA (2013): Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. Clinical Trials, 10(5), 720-734.

Chang W, Cheng J, Allaire JJ, Xie Y, and McPherson J (2018): shiny: Web Application Framework for R. R package version 1.2.0. https://CRAN.R-project.org/package=shiny

11:20 – 11:40 Oliver Sailer[1*], Anna Pöhlmann[1,2], Frank Fleischer[1]

Exploring early interim analyses in basket designs in Oncology

[1] *Boehringer Ingelheim Pharma GmbH & Co. KG*
[2] *University of Bamberg, Bamberg, Germany*

Keywords: Basket design, Bayesian Hierarchical Model, Interim analysis, Oncology

In early clinical development in Oncology, researchers often investigate a single treatment for multiple indications. To make decisions about further development for specific patient cohorts, basket trials are used to evaluate various subpopulations in a single trial for early signs of efficacy. Early futility analyses are performed in these studies to limit the exposure to drug for patients in inactive cohorts. Early interim analyses could also be used to speed up Go decisions for further development.

We perform a simulation study to compare various approaches for interim analyses. Aspects considered include:

- Early futility on cohort level vs. analysis in a Bayesian Hierarchical Model (BHM) allowing for borrowing of information across cohorts
- Early futility only vs. early Go/No-Go decision based on BHM
- Go decision based on BHM with data restricted to cohorts that pass futility vs. Go decision based on BHM with all data including data from futile cohorts.

The BHM used is a modification of the model proposed by Berry et al. (2013). The different design options are evaluated in typical planning scenarios such as all positive, all negative, positive and negative nugget scenarios. Evaluation criteria are false positive and false negative rates as well as study duration and average sample size.

References

Berry SM, Broglio KR, Groshen S, and Berry DA (2013): Bayesian hierarchical modeling of patient subpopulations: efficient designs of phase II oncology clinical trials. Clinical Trials, 10(5), 720-734.

11:40 – 12:00 *Haiyan Zheng, James Wason*

Bayesian basket trial designs for borrowing of information across similar subpopulations

Affiliation: Biostatistics Research Group, Institute of Health and Society, Newcastle University, UK

There has been a recent surge of interest in precision medicine, which goes beyond assessing the population-averaged effects of a new treatment in the conventional paradigms of drug development. A target patient population may be stratified into small subgroups using biomarkers or genetic characteristics. Basket trials provide a framework to evaluate superiority of an experimental therapy over the standard-of-care with respect to the subpopulations. However, considerable criticism has been levelled against designs of early basket trials for the low-powered basketwise analysis strategies.

We propose a novel methodology for phase II basket trials with *k* modules, where information from similar modules can be leveraged to inform inferences in a specific module. For each parameter that underpins the treatment effect in a module, a mixture prior will be specified with the (*k*-1) components being the robust predictive priors obtained from all the other modules, respectively. These (*k*-1) one-to-one predictive priors are robust in that the information from inconsistent modules will be down-weighted by placing a large value on the variance parameter. Hellinger distance between probability distributions of any two module-specific parameters is computed to describe the similarity of information. A symmetric matrix, where the diagonal elements are 0, contains the pairwise Hellinger distance measures and would be updated at the interims of a sequential trial. Each column of this matrix can be converted to a vector of probabilities, serving as the weights attributed to the robust one-to-one predictive priors obtained using information from other modules. This mixture predictive prior will be updated using trial data from the current module to a robust posterior. We illustrate our approach using a PBC trial example. Simulation results will also be presented to show gains such as increase in power, compared with that would have been achieved if information from other modules is discarded.

12:00 – 12:15 Closing Event

& lunch