The International Society for Biopharmaceutical Statistics

The 6th International Symposium on Biopharmaceutical Statistics

Statistical Innovation and Contribution in the Era of Precision healthcare

ABSTRACT BOOK

August 19, 2019

Kyoto International Conference Center, Kyoto Japan
August 26-30, 2019

Table of Contents

Plenary Sessions

Plenary Session I	9:30-10:15AM, August 27, 2019	Sakura, 1F

Adaptive Design of Confirmatory Clinical Trials: Regulatory Perspectives and Recent Advances

Tze Leung Lai, Stanford University

The past decade has witnessed major developments in adaptive design and analysis of clinical trials, from the regulatory guidance publications to industry and from journal publications on methodological advances to address the challenges and opportunities. We focus on late-phase trials that lead to investigational new drug (and device) submissions. In particular, we discuss with examples (a) adaptive subgroup selection in enrichment designs, (b) seamless phase 2-3 designs for cancer therapies, and (c) master protocols and basket trials.

Plenary Session II	8:45-10:15AM, August 28, 2019	Sakura, 1F

Open Science, Data Sharing, and Reproducibility: What's All the Fuss About and Why Should Statisticians Care and Engage in the Journey?

Frank W. Rockhold, Duke University

Calls for greater transparency and 'open data access' in clinical research continue actively. Open access is good for researchers, good for innovation, and most of all good for patients. This is summed up in the Open Science Project (http://openscience.org/) as "If we want open science to flourish, we should raise our expectations to: Work. Finish. Publish. Release."

Open access to individual patient data from clinical trials is an important tool for research in health care. There are certainly challenges and hurdles, but the question should not be whether data should be shared, but rather how and when responsible methods for doing so can be ushered in. Secondary analyses can lead to valuable new insights, enhance the reproducibility and credibility of clinical research, and honor the contributions of trial participants.

This is a difficult task as issues such as patient privacy, academic credit, data standards, and funding are complex challenges that need to be and are being addressed. We need to devise an easier path for clinical trial data sharing and a system that has the full support of sponsors, trialists, secondary researchers, and patients could simplify the process of obtaining data and providing access to secondary researchers. Statisticians and data scientists (both those doing the primary and secondary analyses and involved in providing access) are essential elements of this discussion. This talk will cover the spectrum of challenges and imperatives that are at the core of this critical effort.

Recent Advances in Regulatory Statistics in Imaging Diagnostics and Imaging Precision Medicine

Sue-Jane Wang, US Food and Drug Administration

Plenary Session III	8:45-10:15AM, August 29, 2019	Sakura, 1F

Statistical Innovation and Contribution Today. Will We Need Statisticians Tomorrow?

Robert J. Hemmings, Consilium Salmonson and Hemmings

The introduction of the RCT heralded an era of robust evidence generation for the authorisation of medicines. Statisticians were appointed as champions to optimise design and analysis. Clinical trials increased in size and sensitivity in order to detect small, incremental benefits of new treatments, but have become expensive to conduct and less relevant to clinical practice. Due to an evolution in basic science, both the type of target population and the type of medicinal product has changed. Along with this there has been a revolution in policy. Never has it been more popular to opine on how medicines development should change. Whether due to cost, to science, or to interest in being the architect of revolution, the RCT is under threat, yet there is rarely a viable alternative that can maintain the same rigour. The talk will explore these topics, but will conclude that the R and parts of the C remain fundamental. Part of these considerations relate to the estimand, so the ICH E9(R1) guideline will be described along with another consequence of the ‘era of precision medicine’: sub-setting a broader clinical condition. Throughout this tour of ‘hot topics’ the contribution of statisticians will be explored. We have spent less time than we should in explaining why our principles are important. Perhaps in the future good epidemiology and algorithms for data-mining, AI and machine-learning will suffice? We have an urgent need to improve our communication and influencing skills in a new role as the champions of robust evidence generation.

Regulatory Reform and Challenges of the Japanese Pharmaceutical Regulations

Daisaku Sato, Pharmaceuticals and Medical Devices Agency

In Japan, the regulatory reform has been carried out to improve access to the new therapeutic innovation over the last decade. Presently, the Japanese agency has recorded that their new pharmaceuticals review period is competitive or even faster than the world major agencies.

The 2013 revision of the Pharmaceuticals and Medical Devices Act (PMD Act) stipulated the conditional and time-limited approval scheme for regenerative medical product. So far, 3 products have been granted for the conditional approval. Further, since 2014, “Sakigake” package (as “breakthrough therapy designation” in Japan) has been implemented and 7 products approved.

In 2019, the bill to revise the PMD Act is under pending deliberation at the Diet, which enables early patient access to promising therapies, using Sakigake designation and conditional early approval for pharmaceuticals and medical devices. The Sakigake designation is being legislated in the Act to improve the stability and predictability of the regulatory process.

Pharmaceuticals and Medical Devices Agency (PMDA) has established the Centre for Regulatory Science that is aimed to improve the quality of reviews and consultations, as well as safety measures by promoting regulatory science. PMDA is strengthening the capacity of evidence generation using real-world medical data. It will extend to data in the Medical Information Network System (MID-NET®), which is a hospital-based network of electronic medical records covering more than 4.7 million patients. Along with the regulatory reform, PMDA seeks to improve relevant scientific capabilities and to promote collaboration with all stakeholders to realize its value to human society.

Invited Sessions

Invited Session	10:45AM-12:15PM, August 27, 2019	Sakura, 1F

IS01: Enhancing Regulatory Decision-Making to Support Drug Development: US FDA Plot Programs on Complex Innovative Designs and Model-Informed Drug Development

Promoting the Use of Complex Innovative Trial Designs: An Overview

Dionne L Price, US Food and Drug Administration

The Food and Drug Administration is poised to meet the demands of the changing landscape of drug development as evident by various commitments outlined in the 21st Century Cures Act and the Prescription Drug User Fee Act (PDUFA) reauthorization for fiscal years 2018-2022, known as PDUFA VI. These commitments encompass a broad range of topics and include the advancement and facilitation of the use of complex adaptive, Bayesian, and other novel clinical trial designs. The FDA has initiated several efforts to fulfill the commitments, with the vision of ultimately providing benefit to patients through the development and implementation of innovative designs. FDA's efforts comprise convening a public meeting on promoting the use of such designs, releasing the draft Guidance for Industry: Adaptive Clinical Designs for Clinical Trials and Biologics, and launching the Complex Innovative Trial Design (CID) Pilot Meeting Program. This talk will provide an overview of the FDA's progress on the ongoing CID efforts.

Model-Informed Drug Development at the US Food and Drug Administration: A Perspective on Progress

Issam Zineh, US Food and Drug Administration

Model-informed drug development (MIDD) is a powerful approach to increase efficiency in drug development, impact regulatory outcomes, and improve drug benefit/risk balance. A critical factor for the continued growth of MIDD is a regulatory environment with the capacity, expertise, and transparency that increases stakeholder confidence in MIDD approaches. This talk will provide an update on FDA's MIDD Paired Meeting Pilot Program, which was established to address this need by providing a dedicated avenue for regulatory interaction on MIDD issues.

Invited Session	10:45AM-12:15PM, August 27, 2019	Room 510, 5F

IS03: Recent Development and Challenges in Bioequivalent or Biosimilar Assessment

Statistical Considerations for Demonstration of Analytical Similarity

Harry Yang, Astra Zeneca

Biosimilars are biotherapeutics shown to be of similar quality, efficacy and safety to the reference products. The advances in analytical technologies and particularly in regulatory policies and regulations have powered the biosimilar development, making it one of the fastest-growing sectors. Biosimilar development typically follows a stepwise approach as per regulatory guidelines. Since analytical similarity is the foundation for demonstration of biosimilarity between a proposed biosimilar and a reference product, the process begins with the assessment of critical quality attributes (CQA) used to characterize the biosimilar products. Furthermore, the FDA recommends a tiered system in which quality attributes are categorized into three tiers commensurate with their risk ranking. Different statistical approaches of varying rigor are then used to demonstrate similarity of the tiered quality attributes. These analytical assessments, coupled with PK/PD data and clinical efficacy and safety evaluations, provide the totality of evidence in support of licensure of the product as a biosimilar to the reference product. In this presentation, we discuss statistical considerations for demonstration of analytical similarity.

Recent Development Strategies and Challenges in Biosimilarity Assessment

Heike Woehling, Sandoz Biopharmaceuticals; **Lam Yau**, Sandoz Biopharmaceuticals

Ideally, the clinical development program of a Biosimilar is limited to two clinical studies only: one Pharmacokinetics/Pharmacodynamics (PK/PD) similarity study and one confirmatory efficacy/safety/immunogenicity study.

PK biosimilar studies usually include three treatment arms to demonstrate similar PK between biosimilar vs EU-reference, biosimilar vs US-reference and in addition the PK Bridge EU-reference to US-reference. PK similarity is demonstrated if the 90% confidence interval (CI) for the ratio of geometric means of two products for the PK parameter (typically AUCinf, Cmax and AUClast) fall entirely within the pre-defined margin of 0.80 to 1.25. Various challenges will be addressed with regard to powering the studies adequately for multiple comparisons as well as the possibility to perform an interim blinded re-assessment of the assumed variability to potentially adapt the sample size.

Because efficacy studies are generally not as sensitive as PK/PD studies to identify differences between biosimilar and originator if they exist, the tendency in biosimilar development is towards a PK/PD study only. However, for most products to be developed as a biosimilar, no validated PD biomarker exists.
The objective of a confirmatory efficacy/safety/immunogenicity biosimilar study is to demonstrate that no clinically meaningful differences exist between the products in terms of efficacy, safety and immunogenicity. Examples will be provided for equivalence margins justified from meta-analyses which are the most challenging parts to be agreed upon in advance with various Health Authorities. Additionally challenges with regard to the handling of missing data will be addressed.

Effects of Between-Batch Variability on the Type I Error Rate in Biosimilar Development

Seung-Ho Kang, Yonsei University; **Junhui Park**, Yonsei University

Biological products are known to have some between-batch variation. However, the traditional method to assess biosimilarity does not consider such between-batch variation. Linear random effect models and beta-binomial models are assumed in order to incorporate between-batch variation for the continuous endpoints and the binary endpoints, respectively. Under these assumptions, in this paper we show that the type I error rates are inflated when biosimilarity is evaluated by the traditional method, which ignores between-batch variation.

Statistical Quality Control for Biosimilar Assessment

Hsiao-Hui Tsou, National Health Research Institutes; **Yu-Chieh Cheng**, National Health Research Institutes; **Hsiao-Yu Wu**, National Health Research Institutes; **Ya-Ting Hsu**, National Health Research Institutes

Biosimilars are biological products which are highly similar to the originator biologics in terms of analytical, non-clinical and clinical characterizations. Comparing to expensive original biologics, biosimilars are expected to be more affordable alternative products for patients. However, traditional methods to assess bioequivalence may not be applied to evaluate the similarity between biological products. In this presentation, we will review current methodologies and discuss challenges in biosimilar assessment. We will provide a statistical quality control method for biosimilars. Comparison among all existing methodologies will be demonstrated to explore the better practice in this new and important area.

Invited Session	1:30-3:00PM, August 27, 2019	Sakura, 1F

IS06: Use of Machine Learning and AI for Precision Medicine in Drug Development

Subgroup Identification in Precision Medicine

Xin Huang, AbbVie Inc.; **Yihua Gu**, AbbVie Inc.; **Yan Sun**, AbbVie Inc.; **Ivan S.F. Chan**, AbbVie Inc.

A major challenge in developing precision medicines is the identification and confirmation of patient subgroups where an investigational regimen has a positive benefit-risk balance. In biopharmaceutical development, exploring these patient subgroups of potential interest are usually achieved by constructing decision rules (signature) using single or multiple biomarkers in a data driven fashion, accompanied by rigorous statistical performance evaluation to account for potential overfitting issues inherent in subgroup searching. In this presentation, we will discuss general considerations in exploratory subgroup analysis, investigate statistical methods for subgroup identification, and propose statistical principles for the subgroup performance assessment. An example of subgroup identification for an immunology disease treatment leading to regulatory label inclusion will also be presented.

How to Use Machine Learning Algorithms in Clinical Development

Yoshitake Kitanishi, Shionogi & Co., LTD.; **Masakazu Fujiwara**, Shionogi & Co., LTD.; **Bruce Binkowitz**, Shionogi Inc.

Pharmaceutical Value Chain involves drug discovery, non-clinical, clinical and post-marketing stages, and it is known that machine learning approach can be used to streamline and maximize those stages. In other words, the application concept of machine learning approach can be implemented in each stage, such as finding a correlation structure from big data and making a hypothesis. For example, at the drug discovery stage, classification and prediction of machine learning approach is used for classification of compounds by chemical descriptors and prediction of toxicity, and machine learning approach for hypotheses of drug repositioning across multiple value chains can be used as well.

Recently, there are health insurance claims database and Japanese acute hospital-based claims database as Real World Data after marketing in Japan, and it is considered that classification and prediction of machine learning approach can be applied to them. In other words, the correlation structure of real world is estimated by machine learning. Therefore, we can classify some diseases from various perspectives based on medical histories and patient demographic background, and can predict patient's prognosis for each divided disease by applying classification and prediction methods of machine learning. In this study, we introduce approaches by association analysis and several clustering methods. By applying this approach, pharmaceutical companies may be able to provide appropriate medical care depending on the individual.

The Future Is Now, but Where We Should Focus

Haoda Fu, Eli Lilly & Co.

New technology enables alternative data types, such as unstructured data, streaming data, and big data. There are challenges and opportunities. Naively applying data science package can result in bias and misleading results which will put a risk to patients. What are the common issues? And where we should focus? This talk will provide our perspective based on recent enterprise initiative on advanced analytics and data science from Eli Lilly and Company.

Estimation and Validation of Regressions for Precision Medicine using Real World Data

Lu Tian, Stanford University

While sample sizes in randomized clinical trials are large enough to estimate the average treatment effect well, they are often insufficient for data-driven estimation of treatment-covariate interactions critical to studying precision medicine. Observational data from real world practice may play an important role in alleviating this problem. One common approach in trials is to predict the outcome of interest with separate regression models in each treatment arm, and recommend interventions based on the contrast of the

predicted outcomes. Unfortunately, this simple approach may induce spurious treatment-covariate interaction in observational studies when the regression model is misspecified. Motivated by the need to model the number of relapses in multiple sclerosis patients, where the ratio of relapse rates is a natural choice of the treatment effect, we propose to estimate the treatment effect heterogeneity by coupling the standard regression approach with a doubly robust adjustment that mitigates finding spurious interactions. We also provide a validation procedure for the treatment effect heterogeneity estimator. We conduct simulations to demonstrate the finite sample performance of the proposed methods, and illustrate the advantage of this approach on real data examining the treatment effect of dimethyl fumarate compared to teriflunomide in multiple sclerosis patients.

Invited Session	1:30-3:00PM, August 27, 2019	Room 510, 5F

IS02: Statistical Designs and Considerations in Early Clinical Development

An Adaptive Phase I/II Design

Chin-Fu Hsiao, National Health Research Institutes; **Chieh Chiang**, National Health Research Institutes; **Wen-Jin Guo**, National Health Research Institutes; **Sue-Jang Wang**, US Food and Drug Administration

In recent years, to speed up the development of drug, assessing toxicity and efficacy simultaneously in early phase trials are increasing. Traditional early phase trials mostly assume that both toxicity and efficacy increase monotonically with dose levels. However, due to the immune-biological properties of adoptive cell therapies (ACTs), assuming monotonic does-response for does-finding in early phase trials may be inappropriate. Therefore, Li et al. in 2016 recommended a toxicity and efficacy probability interval (TEPI) design for does-finding in ACT trials. In this study, a modified toxicity and efficacy probability interval design is proposed. It can be seen that our design provides more desirable decisions from safety and efficacy perspectives. An R code is provided for users to use our design to conduct clinical trials, and examine operating characteristics of the designs.

Two-stage Phase I/II Designs

Yuh-Ing Chen, National Central University; **Chen-Wei Yeh**, National Central University

Robust two-stage designs are proposed to identify the optimal biological dose (OBD) of a molecular drug for cancer treatment without assuming any particular model for the dose-toxicity or dose-efficacy relationship. However, toxicity probability is still believed to be increasing with dose and the dose-efficacy relationship may have an ordered pattern, an ascending trend followed by a plateau feature or an umbrella structure with a downturn in the dose range. At the first stage, any model-free method can be used to find an initial estimate of the maximum tolerated dose (MTD) of the drug under study. The second stage of the robust design then takes into account of all the available toxicity and efficacy responses for adaptive MTD estimation and dose assignment to batches of patients. The OBD is finally identified based on the isotonic regression estimates of the efficacy probabilities under the umbrella structure. An extensive simulation is further conducted to investigate how well the proposed two-stage designs can be timely implemented, especially when it takes a long time to observe the efficacy response, and how effectively the robust designs perform on the OBD identification under a variety of dose-toxicity and dose-efficacy relationships.

The Win Ratio: On Interpretation and Handling of Ties

Gaohong Dong, iStats Inc.; **David Hoaglin**, University of Massachusetts Medical School; **Junshan Qiu**, US Food and Drug Administration; **Roland Matsouaka**, Duke University School of Medicine; **Victoria Chang**, BeiGene

To handle composite endpoints, the win ratio has been applied to data analysis and design of clinical trials. Its interpretation, however, is not always clear, and it could handle ties differently. We address these two aspects. First, we express the win ratio as a ratio of two proportions, namely, the proportion of patient level comparisons in which the experimental treatment “wins” over the control divided by the proportion of “wins” for the control, taking into account the priority order of the components. This equivalent form, the ratio of proportions, can ease communication to clinical trial stakeholders—especially when the win proportions themselves are reported. We recommend such presentations. In some simple cases, we connect the win ratio to the odds ratio, the hazard ratio, the Mann Whitney U, and the mean difference. In exploring the role of ties, we introduce the win odds, as an extension of the Mann Whitney odds under the framework of prioritized pairwise comparisons. Finally, we discuss some practical aspects of the win ratio, including rules for defining winners (or losers) and ties, dependence on censoring, win ratio estimands, and benefit risk assessments, as well as applications to two clinical studies.

Statistical Design and Analysis of Rare Diseases Clinical Trials

Shein-Chung Chow, Duke University; **Victoria Chang**, BeiGene

One of the most challenges for rare disease clinical trials is probably the availability of a small patient population. It is then a great concern how to conduct clinical trials with the limited number of subjects available for obtaining substantial evidence regarding effectiveness and safety for approval of the drug product under investigation. FDA, however, does not have intention to create a statutory standard for approval of orphan drugs that is different from the standard for approval of drugs in common conditions. Thus, it is suggested that innovative trial designs such as a complete n-of-1 trial design or an adaptive design should be used for an accurate and reliable assessment of rare disease drug products under investigation. In this article, basic considerations, innovative trial designs, and statistical methods are discussed. In addition, sample size requirements under rare disease setting are derived.

Invited Session	3:30-5:00PM, August 27, 2019	Room 510, 5F

IS14: Innovative Approaches for Trial Design and Analysis

BMA-Mod: A Bayesian Model Averaging Strategy for Determining Dose-Response Relationships in the Presence of Model Uncertainty

A. Lawrence Gould, Merck Sharp & Dohme Corp.

Successful pharmaceutical drug development requires finding correct doses. The issues that conventional dose-response analyses consider, namely whether responses are related to doses, which doses have responses differing from a control dose response, the functional form of a dose-response relationship, and the dose(s) to carry forward, do not need to be addressed simultaneously. Determining if a dose-response relationship exists regardless of its functional form and then identifying a range of doses to study further may be a more efficient strategy. This presentation describes a novel estimation-focused Bayesian approach (BMA-Mod) for carrying out the analyses when the actual dose-response function is unknown. Realizations from Bayesian analyses of linear, generalized linear, and nonlinear regression models that may include random effects and covariates other than dose are optimally combined to produce distributions of important secondary quantities, including test-control differences, predictive distributions of possible outcomes from future trials, and ranges of doses corresponding to target outcomes. The objective is similar to the objective of the hypothesis-testing based MCP-Mod approach, but provides more model and distributional flexibility and does not require testing hypotheses or adjusting for multiple comparisons. Examples illustrate the application of the method.

Practical Considerations of Sequential Analysis of the Restricted Mean Survival Time for Immuno-Oncology Trials

Ying Lu, Stanford University

A progress-free survival (PFS) time is often the primary endpoint in phase 3 oncology trials. The treatment effect of immune-therapies often fails the proportional hazards assumption. A comparison of the restricted mean survival times (RMSTs) has been proposed as an alternative to hazards ratio to evaluate the PFS benefits. One of the challenges for using RMST is to plan sequential clinical trials. Murray and Tsiatis (1999 Biometrics) has presented the statistical formulation for the sequential design for an RMST trial. In this talk, we will discuss some practical considerations, including the timing of interim analysis, confidence intervals of resulted estimated difference in RMSTs, sample size consideration and optimizations. Simulation examples will be presented to illustrate our results.

BOP2: Bayesian Optimal Design for Phase II Clinical Trials with Binary, Co-primary and Other Complex Endpoints

Ying Yan, University of Texas MD Anderson Cancer Center

The endpoints for immunotherapy and targeted therapy are often complicated, making conventional phase II trial designs or commonly used basket designs inefficient and disfunctional. We propose a flexible Bayesian optimal phase II (BOP2) design that is capable of handling simple (e.g., binary) and complicated (e.g., ordinal, nested and co-primary) endpoints under a unified framework. We use a Dirichlet-multinomial model to accommodate different types of endpoints. At each interim, the go/no-go decision is made by evaluating a set of posterior probabilities of the events of interest, which is optimized to maximize power or minimize the number of patients under the null hypothesis. Unlike most existing Bayesian designs, the BOP2 design explicitly controls the type I error rate, thereby bridging the gap between Bayesian designs and frequentist designs. In addition, the stopping boundary of the BOP2 design can be enumerated prior to the onset of the trial. These features make the BOP2 design accessible to a wide range of users and regulatory agencies, and particularly easy to implement in practice. Simulation studies show that the BOP2 design has favorable operating characteristics with higher power and lower risk of incorrectly terminating the trial than some existing Bayesian phase II designs. The software to implement the BOP2 design is freely available at www.trialdesign.org

---

Estimand - An Alternative Implementation of the While on Treatment Strategy

Naitee Ting, Boehringer-Ingelheim Pharmaceuticals, Inc.

Missing data in longitudinal analysis has long been an important concern for drug developers and regulators. ICH published its E9 (R1) guidance in 2017, addressing the issue of estimands. In that guidance, five strategies are introduced. This manuscript focuses on the While on Treatment strategy, and attempts to propose standardized area under the curve (sAUC) as an alternative to LOCF in the application of while on treatment strategy. sAUC is a simple summary statistics reflecting the average change a patient experiences during the trial. It is easy to interpret and easy to understand by non-statisticians. In clinical trials with longitudinal data, sAUC has many applications.

Invited Session	10:45AM-12:15PM, August 28, 2019	Sakura, 1F

IS12: Some Innovative Approaches to Trial Designs and Medical Product Development

Master Protocol Design Considerations in Settings Where Randomized Controlled Trials Are Not Feasible

Sue-Jane Wang, US Food and Drug Administration

Commonly known master protocol has been applied to umbrella, basket and platform trials during the investigational stage of a drug or drugs based on molecular targeting or based on clinical phenotypes. In recent drug developments, the utility of a master protocol has gradually been recognized as a methodologic innovation for various reasons. In this paper, we consider a master protocol design in a post-approval post-marketing requirement setting. Specifically, medical practice after drugs have been approved must be a key element for any proposed novel design to be considered as relevant, practical and feasible for trial implementation. We will give rationales, present a variety of scenarios with examples from Medical Imaging and compare their design efficiency and utilities. Some analytical approaches will also be discussed.

*The views expressed are those of the authors and should not be construed to represent the views or policies of the U.S. Food and Drug Administration.

Sequential, Multiple-Assignment, Randomized Trials for COMparing Personalized Antibiotic StrategieS (SMART-COMPASS)

Scott R. Evans, George Washington University

Patient management is not based on a single decision. Rather, it is dynamic: based on a sequence of decisions, with therapeutic adjustments made over time. Adjustments are personalized: tailored to individual patients as new information becomes available. However, strategies allowing for such adjustments are infrequently studied. Traditional antibiotic trials are often nonpragmatic, comparing drugs for definitive therapy when drug susceptibilities are known. COMparing Personalized Antibiotic StrategieS (COMPASS) is a trial design that compares strategies consistent with clinical practice. Strategies are decision rules that guide empiric and definitive therapy decisions. Sequential, multiple-assignment, randomized (SMART) COMPASS allows evaluation when there are multiple, definitive therapy options. SMART COMPASS is pragmatic, mirroring clinical, antibiotic-treatment decision-making and addressing the most relevant issue for treating patients: identification of the patient-management strategy that optimizes the ultimate patient outcomes. SMART COMPASS is valuable in the setting of antibiotic resistance, when therapeutic adjustments may be necessary due to resistance.

Missing data treatment in Sequential Parallel Comparison Design Studies

Gheorghe Doros, Boston University

Invited Session	10:45AM-12:15PM, August 28, 2019	Room 510, 5F

IS09: Statistical Issues and Methods for Vaccine Development

Potential Study Designs for HIV Vaccine Efficacy Trials in the Era of an Expanding Portfolio of Non-Vaccine HIV Prevention Strategies

Holly Janes, Fred Hutchinson Cancer Research Center

The HIV prevention field has advanced dramatically in recent years through the development of antiretroviral-based prevention strategies such as Treatment-as-Prevention and oral pre-exposure prophylaxis (PrEP). Despite these successes, numerous challenges remain that impede the implementation of these interventions and HIV remains a major global health problem. At the same time, the rapid advancements in the field have complicated the statistical design of efficacy trials of new interventions. For instance, oral PrEP is now part of the standard HIV prevention package offered to trial participants, which poses challenges in ensuring designs have adequate statistical power. In this talk, we discuss potential future efficacy trial designs for evaluating candidate vaccines that accommodate this complex and evolving landscape. We investigate their statistical properties and discuss the challenges that will be faced in seeing these designs implemented, including effective engagement of regulators, ethicists, and community advocates.

Assessment of Correlate of Risk and Protection in Tetravalent Dengue Vaccine Efficacy Trials

Zoe Moodie, Fed Hutchinson Cancer Research Center; **Michal Juraska**, Fed Hutchinson Cancer Research Center; **Ying Huang**, Fed Hutchinson Cancer Research Center and University of Washington; **Yingying Zhuan**, University of Washington; **Youyi Fong**, Fed Hutchinson Cancer Research Center and University of Washington; **Lindsay N. Carpp**, Fed Hutchinson Cancer Research Center; **Steven G. Self**, Fed Hutchinson Cancer Research Center and University of Washington; **Laurent Chambonneau**, Sanofi Pasteur; **Robert Small**, Sanofi Pasteur; **Nicholas Jackson**, Sanofi Pasteur; **Fernando Noriega**, Sanofi Pasteur; **Peter B. Gilbert**, Fed Hutchinson Cancer Research Center and University of Washington; **Jing Jin**, Sanofi R&D Beijing

CYD-TDV is a licensed recombinant live attenuated tetravalent dengue vaccine, given as a 3-dose series on a 0/6/12-month schedule. In two phase 3 trials of the CYD-TDV vaccine performed in Asia (CYD14; 2-14 year olds) and Latin America (CYD15; 9-16 year olds), estimated vaccine efficacy (VE) against symptomatic, virologically confirmed dengue (VCD) occurring between 28 days after the third injection at 12 months (M13) through 25 months was 56.5% and 60.8%, respectively. 50% plaque reduction neutralization test (PRNT50) neutralizing antibody titers to the 4 dengue serotypes were measured at M13 (i.e. one month post-final dose) in a randomly sampled subset of participants and in all VCD cases through month 25. M13 PRNT50 titers were then evaluated for their association with VCD and with the level of VE to prevent VCD based on the VE curve effect modification framework. Participants with higher M13 titer had lower risk of VCD in both the CYD-TDV and placebo groups for each trial and for each serotype, more strongly in vaccine group. Vaccinees with higher M13 average titer to the 4 serotypes had significantly higher VE against VCD of any serotype, indicating that neutralizing antibody titers post-dose 3 correlate with CYD-TDV VE to prevent dengue. High titers associated with high VE for all serotypes and all age groups in each trial, regardless of baseline serostatus. Lowest titers corresponded to zero VE in CYD14 but not CYD15, indicating that other factors besides neutralizing antibodies may influence VE of CYD-TDV (Sanofi Pasteur sponsored and funded the trials).

Quantitative Decision-Making Framework for Phase III Vaccine Efficacy Trial

Wenji Pu, GlaxoSmithKline plc.; **Fabian Tibaldi**, GlaxoSmithKline plc.; **Charlotte Baidoo**, GlaxoSmithKline plc..

A phase III Vaccine Efficacy trial typically requires a very large sample size due to low event rate, and particularly a preventative vaccine efficacy trial demands high efficacy and precision, therefore a super superiority study design is often required. It is very crucial to make decision on moving forward to phase III trial based on the results of a PoC trial. In some vaccine clinical development programs where a PoC vaccine trial has the similar endpoint and population in the phase III, it is much more informative and

appealing to quantify the probability of success in phase III based on the PoC trial results. This talk is to introduce a quantitative decision-making framework by using Bayesian methods for incorporating prior information (such as through eliciting expert opinion) into the study design and planning for a vaccine clinical development program, and setting Go/No-go criteria and using Bayesian predictive probability of success in phase III to calibrate the PoC trial decision rules.

Statistical Challenges of an Immunobridging Approach to Assess Clinical Benefit as the Basis for Licensure of a Prophylactic Ebola Vaccine

Bart Spiessens, Janssen Research & Development; **Thierry Van Effelterre**, Janssen Pharmaceutica N.V.; **Jan Serroyen**, Janssen Pharmaceutica N.V.; **Viki Bockstal**, Janssen Vaccines & Prevention; **Ramon Roozendaal**, Janssen Vaccines & Prevention; **Cynthia Robinson**, Janssen Vaccines & Prevention; **Kerstin Luhn**, Janssen Vaccines & Prevention; **Jenny Hendriks**, Janssen Vaccines & Prevention; **Benoit Callendret**, Janssen Vaccines & Prevention; **An Vandebosch**, Janssen Pharmaceutica N.V.

Janssen is developing an Ebola vaccine using a 2-dose heterologous regimen with Ad26.ZEBOV vaccination followed by MVA-BN-Filo in a 56 days interval. In the absence of the possibility to demonstrate clinical efficacy data, an experimental approach has been discussed with regulatory authorities that would be considered appropriate to support licensure. Principle agreement was reached that bridging immunogenicity and efficacy data obtained in non-human primates (NHP) to clinical immunogenicity data in humans could provide a reasonable likelihood that the vaccine regimen will provide clinical benefit and serve as the basis to support the Animal Rule (FDA) / Conditional Approval (EMA) licensure pathways.

To achieve this a statistical immunobridging model was developed using NHP challenge studies with survival/death as outcome and the binding antibody concentration assessed by the EBOV GP Filovirus Animal Nonclinical Group (FANG) enzyme-linked immunosorbent assay (ELISA) as predictor. Clinical benefit in humans will be assessed by calculating a mean predicted survival probability based on this logistic regression model and by using the antibody concentrations in humans measured in the same ELISA from the Phase 2 and 3 clinical studies. The 95% CI confidence interval on the mean predicted survival probability will be computed with the double bootstrap method. This approach is similar to the bridging approach as described by Schiffer et al (2015) for the anthrax vaccine. This presentation will focus on the statistical and regulatory challenges that were encountered during the research that was performed and ultimately led to the development of the proposed strategy.

Invited Session	1:30-3:00PM, August 28, 2019	Sakura, 1F

IS04: Rethinking Estimators within the Estimand Framework

Semiparametric Copula-based Analysis for Treatment Effects in the Presence of Treatment Switching

Chia-Hui Huang, National Taipei University; **Yi-Hau Chen**, Institute of Statistical Science, Academia Sinica; **Jinn-Li Wang**, Taipei Medical University; **Mey Wang**, Taiwan Center for Drug Evaluation

In controlled trials, “treatment switching” occurs when patients in one treatment group switch to the alternative treatment during the trial, and poses challenges to evaluation of the treatment effects owing to crossover of the treatments groups. In this work, we assume that treatment switches occur after some disease progression event, and view the progression and death events as two semicompeting risks. The proposed model consists of a copula model for the joint distribution of time-to-progression (TTP) and overall survival (OS) before the earlier of the two events, as well as a conditional hazard model for OS subsequent to progression. The copula model facilitates assessing the marginal distributions of TTP and OS separately from the association between the two events, and, in particular, the treatment effects on TTP and on OS in the absence of treatment switching. The proposed conditional hazard model for death subsequent to progression allows us to assess the treatment switching (crossover) effect on OS given occurrence of progression and covariates. General semiparametric transformation models are employed in the marginal models for TTP and OS. A nonparametric maximum likelihood procedure is developed for model inference, which is verified through asymptotic theory and simulation studies. The proposed analysis is applied to a lung cancer dataset to illustrate its real utility.

What Estimands Do Recurrent Event Data Approaches Estimate When Terminal Event Exists?

Jiawei Wei, Novartis Pharma AG; **Mouna Akacha**, Novartis Pharma AG; **Tobias Mütze**, Novartis Pharma AG; **Heinz Schmidli**, Novartis Pharma AG; **Dong Xi**, Novartis Pharma AG

Recurrent event data are multivariate failure time data where the individuals experience repeated occurrences of the same type of event. Many generalizations of the Cox proportional hazard method have been proposed to analyze recurrent event data. In clinical trials where recurrent events are clinically meaningful and where treatments are expected to impact the first as well as subsequent events, the Negative Binomial (NB) model and the Anderson-Gill model with a robust variance estimator (LWYY) are the most commonly used methods to evaluate the treatment effect. The recent ICH (2017) guideline emphasizes that trial objectives and statistical approaches (estimators) should be aligned by clearly defining the estimand of interest. Under the estimand framework, we focus on the setting where a terminal intercurrent event exists (e.g. death), and define two estimands with causal interpretation. Then, we investigate the parameters targeted by NB and LWYY models in the presence of an associated terminal event, to better understand whether they align with the estimands we defined, both numerically and by simulation. Finally, we apply the two estimands to a case study on chronic heart failure, and appropriate estimators are applied to analyze the data.

Mixture of Multivariate t Linear Mixed Models with Missing Information

Tzy-Chy Lin, Taiwan Center for Drug Evaluation

Linear mixed-effects (LME) models have been widely used for longitudinal data analysis as it can account for both fixed and random effects, while simultaneously incorporating the variation on both within and between subjects. In clinical trials, some drugs may be more effective in Westerners than the Orientals. In this situation, such heterogeneity can be modeled by a finite mixture of LME models. The classical modeling approach for random effects and the errors parts are assumed to follow the normal distribution. However, normal distribution is sensitive to outliers and intolerance of outliers may greatly affect the model estimation and inference.

In this presentation, we propose a robust approach called the mixture of multivariate t LME models with missing information. To facilitate the computation and simplify the theoretical derivation, two auxiliary permutation matrices are incorporated into the model for the determination of observed and missing components of each observation. We describe a flexible hierarchical representation of the considered

model and develop an efficient Expectation-Conditional Maximization Either (ECME) algorithm for carrying out maximum likelihood estimation. Simulation results and real data analysis are provided to illustrate the performance of the proposed methodology.

Estimands and Estimation of Population-Averaged Parameters in Randomized Clinical Trials

Tosiya Sato, Kyoto University School of Public Health

In the step 2 document of the ICH E9.R1 “Estimands and Sensitivity Analysis in Clinical Trials,” there are four attributes of estimands, namely, the population, the variable (or endpoint), the specification of how to account for intercurrent events, and the population-level summary of the variable. Comparisons of simple summary measures, such as proportions and means, between treatment groups give the population-level summary of treatment effects in a perfectly conducted trial, in which covariates are well balanced, the trial participants are well complied their assigned treatment and, there are no loss to follow-up. It is because randomization allows us to estimate a counterfactual expectation, E (Endpoint | set Treatment=t), by an observed expectation, E(Endpoint | Treatment=t), where, for example, t=1 for the test treated and t=0 for the control treated. And hence, inference for treatment effects should not be based on random sampling framework given in typical statistics textbooks, but on randomization framework.

When there are some irregularities, estimation of population-averaged treatment effects is not straightforward. In the case of covariate imbalance, one can conduct an adjustment of such an imbalance via regression models. However, we cannot interpret an estimated regression coefficient as a population-averaged treatment effect for non-linear, e.g., logistic, models and/or liner or log-liner models with interaction between treatment and covariates.

We will discuss the estimation of population-averaged treatment effects using the model-based standardization, and extend to apply doubly robust estimation procedures. We will also discuss randomization-based analyses of averaged treatment effects under intercurrent events.

Invited Session	1:30-3:00PM, August 28, 2019	Room 510, 5F

IS07: Adaptive Designs for Small Population Clinical Trials

A Bayesian Nonparametric Utility-Based Design for Comparing Treatments to Resolve Air Leaks After Lung Surgery

Peter Müller, University of Texas at Austin

We propose a Bayesian nonparametric utility-based group sequential design for a randomized clinical trial to compare a gel sealant to standard care for resolving air leaks after pulmonary resection. Clinically, resolving air leaks in the days soon after surgery is highly important, since longer resolution time produces undesirable complications that require extended hospitalization. The problem of comparing treatments is complicated by the fact that the resolution time distributions are skewed and multi-modal, so using means is misleading. We address these challenges by assuming Bayesian nonparametric probability models for the resolution time distributions and basing the comparative test on weighted means. The weights are elicited as clinical utilities of the resolution times. The proposed design uses posterior expected utilities as group sequential test criteria. The procedure's frequentist properties are studied by computer simulation. If time permits we will also briefly discuss another application of BNP to comparing treatments in the presence of semi-competing risks. https://www.ncbi.nlm.nih.gov/pubmed/28959372

Clinical Trial Designs with Data-Driven Selection of Subgroups

Franz König, Medical University of Vienna

Increasing Evidence Designs in Small Population Clinical Trials

Andreas Faldum, University of Münster

Proof of evidence towards or against a new approach takes a long time with conventional study designs in small population clinical trials.

This talk presents a design that increases evidence stepwise. Thereby, subsequent stages in an adaptive-sequential design with sharpening futility bounds realize increasing evidence. At the first stage, the procedure starts with a liberal futility bound which allows dropping a very poor performing treatment approach. This liberal futility bound yields a high power to prevent an erroneous futility stop with a comparatively small sample size. Each next stage iteratively raises the hurdle for a new treatment by decreasing its respective futility bound.

Firstly, the approach enables the control of the rejection-for-futility rate i.e. the trial stops in case of a poor performing new therapy by construction. Moreover, the uninflated α level controls the rejection of the null hypothesis at interim and final analyses. Sample size can be adapted after each interim analysis.

If the trial has to stop i.e. for financial reasons or since a new promising treatment is forthcoming, one minus the smallest futility bound successfully passed can be interpreted as the level of evidence attained in the trial. This interpretation broadens the concept of a level α test.

In view of an increasing number of therapy approaches due to individualized medicine, the increasing evidence design concentrates the resources in those approaches that appear most promising.

The work is part of “Adaptive Designs in Individualized Therapy (ADIT)” supported by the Federal Ministry of Education and Research (FKZ 01EK1503A).

Experiences with Adaptive Design Clinical Trials for Medical Device Development

Toshimitsu Hamasaki, National Cerebral and Cardiovascular Center; **Haruko Yamamoto**, National Cerebral and Cardiovascular Center; **Mayumi Fukuda**, National Cerebral and Cardiovascular Center

Due to the size, availability and accessibility of the population to be studied, we often face challenges in designing clinical trials for medical devices with enough number of participants to answer the primary research question. The resulting need for new approaches to efficient trial designs which aim to assess medical devices in a risk-benefit, rigorously controlled manner, in small populations to gain the most information from the available data.

In this talk, we will describe our recent two experiences in designing medical device clinical trials for regulatory submission, using adaptive designs; one is the extracorporeal continuous-flow ventricular assist device in patients with severe heart failure or refractory cardiogenic shock, and the other is a resorbable surgical sealant device vascular reconstructions to achieve adjunctive hemostasis in patients with stable pump or extra-corporeal membrane oxygenation surgical operation. We will discuss statistical issues in designing the clinical trials and developing the trial protocols, and share communications with Pharmaceuticals and Medical Devices Agency.

Invited Session	3:30-5:00PM, August 28, 2019	Sakura, 1F

IS18: Inference and Decision Making for Contemporary Drug Development and Approval

Journey of Bayesian Inference and Decision Making on Drug Development and Approval

J. Jack Lee, University of Texas MD Anderson Cancer Center

Statistical analysis and inference form a bedrock of providing sound evidence for drug approval. The P-value based frequentist inference has been the standard for clinical trials. A trial is considered successful if the P-value is less than or equal to 0.05 and failed otherwise. This null hypothesis significant testing approach has notable limitations, which could account for the failures of many drugs in clinical development. In contrast, Bayesian approach provides a natural and consistent framework for statistical inference and decision making. It can synthesize evidence by incorporating prior distribution and data to form the posterior distribution to directly estimate the parameter of interest. Information acquired from other trials can be incorporated into the prior distribution to increase the estimation efficiency. Bayesian hierarchical model allows borrowing information across subgroups. Bayes factor gives an appealing alternative to P-value in testing hypothesis. Applying the decision theoretical approach, informed decision can be made by constructing proper utilities to address difficult questions such as the efficacy/toxicity tradeoff and the cost/benefit analysis in public health policy making. Examples of drug development and approval applying Bayesian methodology will be discussed.

Update Your Prior: Use of Bayesian Methods in Drug Development

Martin Posch, Medical University of Vienna; **Nicolas Ballarini**, Medical University of Vienna; **Franz König**, Medical University of Vienna

The formal regulatory assessment of experimental evidence from clinical trials is typically based on frequentist approaches. However, in difficult experimental settings as the development of orphan medicines or the investigation of medicines in vulnerable populations, current regulatory guidance highlights the potential of Bayesian methods for decision making [1]. An example is drug development in pediatric populations, where large clinical trials may not be feasible or ethical. For this setting, an extrapolation framework has been proposed that uses Bayesian tools to incorporate evidence generated in the adult population [2].

While regulatory acceptance of the Bayesian paradigm for the analysis of confirmatory trials is currently limited to special experimental settings, these methods can bring more generally a substantial benefit for the design of frequentist clinical trials. Especially, Bayesian decision theoretic approaches have been proposed to optimize complex clinical trials, as adaptive designs, biomarker guided trials and multi-arm, multi-stage designs.

[1] Guideline on clinical trials in small populations, 2006, CHMP/EWP/83561/2005

[2] Reflection paper on the use of extrapolation in the development of medicines for paediatrics, 2018, EMA/189724/2018

Invited Session	3:30-5:00PM, August 28, 2019	Room 510, 5F

IS10: Recent Development on Missing Data Issues under ICH E9 (R1) Estimand Framework

Principal Stratification: A Strategy for Intercurrent Events that Lead to Unascertainable Outcomes or Confounding

Bohdana Ratitch, Eli Lilly & Co; **Ilya Lipkovich**, Eli Lilly & Co

Certain post-randomization intercurrent events (ICEs) make it impossible to perform planned assessments of clinical outcomes for some subjects in clinical trials. For example, it is impossible to evaluate Health Related Quality of Life (HRQoL) in subjects who die before the planned assessment time. Other ICEs may represent changes in study treatment that introduce confounding into the inference about causal treatment effects. An example is an initiation of rescue therapy when the objective is to compare initial treatments without the effect of rescue. In some cases involving these types of ICEs, principal stratification strategy may be an appropriate tool to define clinically meaningful estimands. This is the case when it is of interest to estimate a treatment effect in a subject population ("stratum") whose status with respect to the ICE would be identical, irrespective of treatment group. For example, an objective may be to compare the effects of two treatments on HRQoL or a laboratory parameter after a certain treatment period in subjects who would survive to the timepoint of interest on either treatment. In a parallel-group design, it is not possible to directly observe which stratum the subject belongs to because we observe what happens only on the treatment to which the subject was randomized. Membership of a principal stratum must therefore be inferred from pre-randomization covariates, typically using statistical modelling. In this presentation, we will discuss some examples where principal stratification can be of interest, the definition of estimands using this strategy, and some methods for principal stratification analysis.

Post E9 (R1) World: Points to Consider from Industry's Point of View

Satoru Tsuchiya, Sumitomo Dainippon Pharma, Co., Ltd.

ICH E9 (R1) guideline has been discussed since 2014, and it becomes finally available soon. "A mindset" is an anagram of "estimand"; so I will talk about the mindset after E9 (R1) world from industry points of view.

With the basis of concept of estimand framework, there are some points to consider on planning of clinical trial, describing study protocol, collecting data, reporting study results, etc.

This guideline does not say that any particular estimand is always great universally. Rather, I believe it states the importance of thinking the estimand for each clinical trial carefully at planning stage. In other words, this guideline provides a good opportunity to consider what is the treatment effect to be obtained in clinical trial. Describing clear treatment effect is very useful for all stakeholders including patients.

Implementation of Estimand Concept in Immunology Disease Area

Na Hu, Boehringer-Ingelheim Inc.

The presentation will discuss the choice of primary estimands, considering various intercurrent events, especially the common use of rescue medications in the immunology disease area. We will also cover the implications of estimand concept implementation on trial design, protocol language, trial conduct and statistical analyses. The focus will be on binary data together with statistical methods on handling missing data. A hypothetical example will be used for illustration.

On Statistical Methods for Some Common Hypothetical Estimands in Clinical Trials

Frank Liu, Merck Sharp & Dohme Corp.

Hypothetical estimands are one category of estimands defined in the draft ICH E9 (R1) addendum. Within the hypothetical estimand framework, the treatment and outcome after intercurrent events are envisioned under some assumed conditions. This provides more flexibility for sponsors and regulatory agencies to

assess treatment effect under different anticipated scenarios for patients who discontinue study medication early. With different assumptions, several estimators may be used to evaluate the treatment effect for a given hypothetical estimand. In this presentation, we will discuss a few common hypothetical estimands in longitudinal clinical trials where the primary analysis is at a given time point, and assumptions are required to handle missing data for patients who discontinue from the study medication prior to the analysis time point. Commonly used methods such as mixed effect models for repeated measures (MMRM), control-based imputation (copy-reference, jump-to-reference, copy increment of reference), and return-to-baseline will be included. Statistical properties for likelihood- and multiple-imputation- based approaches will be explored. The methods will be illustrated with real clinical trial datasets.

Invited Session	10:45AM-12:15PM, August 29, 2019	Sakura, 1F

IS05: Demystifying Estimands for Life History Processes

Recurrent Event Analysis Yielding Estimands with a Causal Interpretation

Richard J. Cook, University of Waterloo

In studies of chronic diseases individuals are often observed to experience recurrent adverse events and the goal in many clinical trials is to reduce the risk of their occurrence. Despite the use of randomization when allocating the treatment at study entry, causal inference is challenging when a) individuals may be lost to follow-up or withdrawn from a study at a dependent censoring time, or b) the recurrent event process ends upon the occurrence of a terminal event such as death. The issues in these settings will be described with specific examples, and alternative estimands discussed to address the limitations of common hazard or rate based estimands.

Assessment of a Treatment Effect for Recurrent Event Data in the Presence of a Terminal Event

Philip Hougaard, Lundbeck

The paper considers clinical trials where multiple occurrences of the same event are recorded over time. The frame is events that can occur at most a few times during a trial, such as hospitalizations or heart failures. I will first consider the case without terminal events (meaning not considering the possibility of death), focusing on, the Poisson and the frailty-Poisson model, with a proportional hazards assumption. Multi-state models will also be mentioned but are less convenient for this purpose. The frailty-Poisson model shows the same treatment effect unconditionally as conditionally on the frailty. From this we will learn that studying the first event only is insufficient because it suffers from a selection effect implying that the unconditional effect is smaller than the conditional. Terminal events, such as death, make the case more complicated because events cannot occur after death and therefore we need to be concerned whether any suggested analysis technique could make a treatment with high mortality appear as successful in reducing the number of events. The suggestion is to consider the integrated hazard of events. This will be discussed in perspective of the estimand definition. Estimands is a relatively new concept that has entered guidelines for the statistics in the pharmaceutical industry. Basically, the concept is a formalization of the handling of missing data.

Efficiency Comparison of Time-To-First and Recurrent Event Analyses with A Focus on Heart Failure Trials

Arno Fritsch, Bayer

In the past, the standard (composite) primary endpoint in heart failure (HF) trials was the time to either HF hospitalization or cardiovascular death, whichever occurs first. With improved treatments on the market HF has now changed from a short-term and quickly fatal condition to a chronic disease, characterized by recurrent HF hospitalizations and high mortality. Therefore, there is interest to use a primary analysis based on recurrent events rather than only the first event. This would better characterize and quantify the full disease burden of HF and one expects to achieve a power increase due to incorporating more statistical information.

In an extensive simulation study we have compared standard methods for analyzing recurrent event data with time-to-first event analyses in the setting of a typical heart failure trial. Methods included the negative binomial model and the Andersen-Gill model with robust standard errors. As an example of rank-based methods, the WinRatio approach was also investigated. Scenarios where the treatment effect on HF hospitalizations differed from that on CV death were a special focus of the simulation study. We will present the results of this simulation study and also consider related estimands within the draft ICH E9 addendum framework.

Causal Mediation of Semi-competing Risks

Yen-Tsung Huang, Institute of Statistical Science, Academia Sinica

The semi-competing risk problem arises when one is interested in the effect of an exposure or treatment on both intermediate (e.g., having cancer) and primary events (e.g., death) where the intermediate event may be censored by the primary event, but not vice versa. Here we propose a nonparametric approach casting the semi-competing risks problem in the framework of causal mediation modeling. We set up a mediation model with the intermediate and primary events, respectively as the mediator and the outcome, and define indirect effect (IE) as the effect of the exposure on the primary event mediated by the intermediate event and direct effect (DE) as that not mediated by the intermediate event. A Nelson-Aalen type of estimator with time-varying weights is proposed for direct and indirect effects where the counting process at time t of the primary event $N_{2n_1}(t)$ and its compensator $A_{n_1}(t)$ are both defined conditional on the status of the intermediated event right before t, $n_1(t^-) = n_1$.We show that $N_{2n_1}(t) - A_{n_1}(t)$ is a zero-mean martingale. Based on this, we further establish the asymptotic unbiasedness, consistency and asymptotic normality for the proposed estimators. Numerical studies including simulation and data application are presented to illustrate the finite sample performance and utility of the proposed method.

Invited Session	10:45AM-12:15PM, August 29, 2019	Room 554, 5F

IS15: Issues in Adaptive and Complex Clinical Trials Designs

Patient-Centred Clinical Trial Designs to Support Precision Healthcare

Andrew P. Grieve, UCB Pharma

Glioblastoma is the most common and aggressive malignant brain tumour in adults, with a poor prognosis and tumours tend to recur after standard multi-modal treatment. Under such circumstances patients may be willing to face a higher than 2.5% chance of being treated with an ineffective drug given that there are no effective alternatives. Recent work by decision analysts and statisticians from the CDRH division of the FDA has investigated how to incorporate patient preferences into the regulatory approval process based on Bayesian Decision Analysis leading to a dependence of decision thresholds on disease severity. These developments are like ideas that have been looked at by pharmaceutical statisticians. In this talk we review the latest developments in the context of patient-centred clinical trials. These developments are in the spirit of Don Berry's observation that "We should also focus on patient values, not just p-values.

Bias and Type I Error of Promising Zone Designs Testing One or More Hypotheses

Florian Klinglmüeller, Austrian Medicines & Medical Devices Agency; **Franz König**, Medical University Vienna

Promising zone designs permit to increase the sample size based on unblinded interim data, but do not require to adjust the statistical test or significance level in order to control the overall Type I error rate. Recent articles have extensively studied the operating characteristics of such designs in terms of power and corresponding sample size requirements. We investigate the bias of estimates using the unweighted mean statistic and the impact on the multiple Type I Error rate when hierarchically testing primary and secondary endpoints following a promising zone design.

We evaluate the bias following different sample size rules (predictive power rules, conditional power rules, rules that maximize the bias) and the Family Wise Error Rate for testing strategies where the secondary endpoint is tested using the conventional pooled test statistic, if the primary endpoint is rejected.

Adapting Study Designs Based on All Available Information from Baseline up to the Primary Endpoint: Is It worth the Effort?

An Vandebosch, Janssen Research and Development; **Kelly Van Lancker**, Ghent University; **Stijn Vansteelandt**, Ghent University and London School of Hygiene and Tropical Medicine

In adaptive designs, conditional power is a frequently used statistical tool to guide decisions on sample size modifications or potential stopping for futility. This is typically based on the subset of subjects who have reached the primary endpoint for evaluation. However, the available information captured in baseline covariates, or short-term endpoints can improve the precision of treatment effect estimates when they are predictive for the primary endpoint. For example, in patients diagnosed with severe infectious diseases, the clinical condition of infection is often evaluated at several timepoints during treatment while the primary endpoint evaluation of cure for a subject is only assessable after a sufficiently long treatment-free evaluation period.

Based on a motivating example, this work presents an interim decision procedure based on conditional power that allows stopping for futility and/or (unblinded) sample size modifications while employing all available information: available primary endpoint data as well as baseline covariates and on-treatment read-outs of subjects recruited in the trial however not reached the primary timepoint for evaluation yet.

The proposed procedure can enable efficiency gains: increased power of the decision procedure as well as reductions in expected sample sizes without compromising the Type I error of the primary analysis through the p-value combination test. Furthermore, this procedure is robust against misspecification of the adopted prediction models. This will be illustrated through simulation studies for a variety of scenarios based on data from the motivating case study.

From Adaptive Designs to Complex Innovative Trials: What Has Changed?

Yannis Jemiai, Cytel

A rose by any other name would smell as sweet," wrote Shakespeare. Are complex innovative designs (CID) mere rebranding of adaptive designs? It would seem so, yet perhaps this is warranted. Despite decades of contributions to statistical methodology and application, encouragement by industry working groups and regulatory authority guidance documents, adaptive and more importantly optimal study designs remain seemingly underutilized in the development pathway of new medicines. The reasons for this may be many, including slow dissemination of knowledge, inadequate access to software tools, and regulatory or operational hurdles. However, for the most part, innovation is stymied by a fear of leading in the absence of good examples, and an inability for statisticians to express the business value of alternative design options. In designating certain study designs as complex innovative designs, we have de facto rendered certain kinds of adaptive designs acceptable or accepted. Yet together with the introduction of other concepts, such as estimands and risk-based monitoring, adaptive and complex innovative designs are merely a way to make optimal decisions and realize business value for stakeholders.

When and How Should Precision Medicine Trials Be Adaptive?

James Wason, Newcastle University and University of Cambridge

There is an increasing interest in designing clinical trials to enable precision medicine. Precision (also known as stratified or personalised) medicine is about going beyond whether a treatment works on average to finding which patient groups it works particularly well or poorly for.

The traditional randomised controlled trial can provide good information through subgroup analysis. However by prospectively utilising patient subgroups, which are thought to be associated with treatment effect, in the trial design we may improve the power and patient benefit provided by the trial.

There are many approaches to doing this, using adaptive and non-adaptive approaches. In this talk I will cover a range approaches that have been proposed and discuss some considerations for whether an adaptive design is useful. I will cover: 1) Basket trials, where a treatment is tested in multiple subgroups or related conditions with adaptive borrowing of information possible; 2) Umbrella trials, where multiple treatments are tested within multiple patient subgroups, often with an adaptive design to best match treatments with groups; 3) adaptive signature design, where a large number of potential biomarkers (e.g. genomic information) can be used to form a signature predicting patients who receive a lot of benefit from a new treatment.

I will aim to cover recent work and lay out future research needed to realise their potential.

Invited Session	1:30-3:00PM, August 29, 2019	Sakura, 1F

IS16: Oncology trials with Non-Proportional Hazards

A Modestly Weighted Logrank Test

Carl-Fredrik Burman, AstraZeneca R&D Gothenburg; **Dominic Magirr**, Cambridge Cancer Genomics

We propose a new class of weighted logrank tests (WLRT) that control the risk of concluding that a new drug is more efficacious than standard of care, when, in fact, it is uniformly inferior. Perhaps surprisingly, this risk is not controlled for WLRT in general. Tests from this new class can be constructed to have high power under a delayed-onset treatment effect scenario, as well as being almost as efficient as the standard logrank test under proportional hazards.

Design Challenges in the Era beyond Proportional Hazard Assumptions

Armin Schueler, Merck KGaA;

For the planning and analysis of time-to-event endpoints proportional hazards (PH) are often assumed. As a consequence, the logrank test (optimal under PH) and hazard ratios are used as standard tools for the analysis. However, delayed treatment effects, disease progression, treatment switchers or the presence of subgroups with differential treatment effects may challenge the assumption of proportional hazards.

This presentation will evaluate the simultaneous impact of different sources of non-proportional hazards on the survival and hazards functions. In addition, the operating characteristics of the logrank test and potential alternatives will be assessed.

The Consideration of Non-Proportional Hazards when Choosing a Randomization Procedure in Survival Studies

Viviane Rückbeil, RWTH Aachen University; **Ralf-Dieter Hilgers**, RWTH Aachen University

The susceptibility of a study to allocation bias is related to the randomization procedure used to allocate patients to treatments. Consequently, the choice of a suitable randomization procedure is a crucial part of the study design. Based on the unweighted log rank test and assuming proportional hazards, a study-specific evaluation model for the selection of a suitable randomization procedure has been developed. However, the assumption of proportional hazards is often not given in practice due to changing treatment effects over time. To increase statistical power in the situation of non-proportional hazards, a weighted version of the log rank test is very often considered.

We propose a semiparametric bias model to reflect the presence of selection bias in a two-arm parallel group trial with a survival outcome. Based on this model, we present an approximation to quantify the impact of selection bias on the rejection probability if the treatments are compared using a (weighted) log rank test. The selection of a suitable randomization procedure with respect to the expected type I error probability is illustrated using the unweighted and weighted log rank tests.

We demonstrate that the evaluation results of the weighted and unweighted log rank tests differ, but that usually the same randomization procedure tends to perform best. If the weights are known in advance, the selection should be made using the weighted log rank test. Otherwise, the unweighted log rank test can also be used to select a suitable randomization procedure.

Evaluating the Impact of Delayed Effects in Oncology Confirmatory Clinical Trials

Jose Luis Jimenez, Novartis Pharma AG; **Viktoriya Stalbovskaya**, Merus; **Byron Jones**, Novartis Pharma AG

Proportional hazards are a common assumption when designing confirmatory clinical trials in oncology. This assumption not only affects the analysis part but also the sample size calculation. The presence of

delayed effects causes a change in the hazard ratio while the trial is ongoing since at the beginning we do not observe any difference between treatment arms and after some unknown time point, the differences between treatment arms will start to appear. Hence, the proportional hazards assumption no longer holds and both sample size calculation and analysis methods to be used should be reconsidered. The weighted log-rank test allows a weighting for early, middle and late differences through the Fleming and Harrington class of weights, and is proven to be more efficient when the proportional hazards assumption does not hold. The Fleming and Harrington class of weights, along with the estimated delay, can be incorporated into the sample size calculation in order to maintain the desired power once the treatment arm differences start to appear. In this article, we explore the impact of delayed effects in group sequential and adaptive group sequential designs, and make an empirical evaluation in terms of power and type-I error rate of the of the weighted log-rank test in a simulated scenario with fixed values of the Fleming and Harrington class of weights. We also give some practical recommendations regarding which methodology should be used in the presence of delayed effects depending on certain characteristics of the trial.

Invited Session	1:30-3:00PM, August 29, 2019	Room 510, 5F

IS13: Statistical Methods in Drug Development

Accounting for Pilot Study Uncertainty in Sample Size Determination of Randomized Controlled Trials

Yaru Shi, Merck Sharp & Dohme Corp.; **Fang Li**, Merck Sharp & Dohme Corp.; **Se Li**, Merck Sharp & Dohme Corp; **Jie Chen**, Merck Sharp & Dohme Corp.

Appropriately sized randomized controlled trials not only provide precise estimates of treatment effects for planning future randomized controlled trials (RCT), but also potentially reduce the total sample sizes for both pilot studies and RCTs. This paper summarizes some existing methods and proposes a new tolerance probability based approach for taking into account the pilot study variability in the sample size determination of future RCTs. These methods are compared in terms of the total sample sizes for both pilot studies and RCTs. Formulas are presented based on normally distributed responses and further developments are given in examples with binary and time-to-event outcome outcomes. Simulations are performed to compare these approaches in terms of total sample sizes given a set of predefined criteria. Some further discussions are presented and recommendations are given for different designs of the pilot study and the RCT.

When Convention Meets Practicality: Pooled Analysis Testing under the Two-Study Paradigm

Frank Bretz, Novartis; **Dong Xi**, Novartis; **Willi Maurer**, Novartis

Standard considerations on multiplicity challenges arising within a single clinical study include comparing several doses of a new drug for more than one outcome variable. A different source of multiplicity arises when evidence is collected across multiple studies. For example, it is common practice to request two statistically significant confirmatory Phase III clinical studies (so-called two-study paradigm), which provides a stringent replicability standard in pharmaceutical drug development. Phase III studies are usually designed with a large sample size to provide sufficient exposure of the new drug in the target population. As a result, managing two large long-term studies may pose statistical, logistical and practical challenges. Motivated by a case study, we discuss in this presentation alternative test strategies that control the error rate across studies at an appropriate significance level while maintaining study-level familywise error rate conventionally. The proposed approaches are simple to communicate with clinical teams and other stakeholders.

Roles of Frailty in Modelling Competing-Risks Data: Assessing Treatment Effect

Il Do Ha, Pukyong National University

Frailty, the effect of unobserved random covariates on the risk of a patient, is very useful in modelling heterogeneity and/or dependence among time-to-event data. Until now, frailty models with unspecified baseline hazard have been widely used for the analysis of various survival data; particularly, ignoring this frailty can lead to biased estimate of treatment effect. However, for the model inference the frailty term has been often regarded as a nuisance, leading that it does not provide directly inference on frailty due to elimination of frailty term by integration. In this talk, we introduce various roles of frailty in analyzing competing risks (CR) data allowing for different types of events which can be correlated, via h-likelihood (Lee and Nelder, 1996; Ha et al., 2017). Unlike the classical likelihood for fixed parameters only, the h-likelihood is constructed for both fixed parameters and unobserved frailties at the same time, and it provides efficient statistical inference for various univariate and multivariate survival models, together with **frailtyHL** R package (Ha et al., 2017). For the purpose, we consider to add frailty term into two different but popular CR models, i.e. cause-specific and sub-distribution hazard models. In particular, with inference of individual frailties in both CR models we show how to investigate the heterogeneity of treatment effect across centers (i.e. random treatment-by-center interaction) in multi-center clinical trial under competing risks setting. Furthermore, we also present the usefulness of frailty in modelling semi-competing risks data where only a terminal event (e.g. death) censors a non-terminal event (e.g. disease recurrence); a patient may experience both events that may be correlated.

Lee, Y. and Nelder, J. A. (1996). Hierarchical generalized linear models (with discussion). Journal of the Royal Statistical Society B, 58, 619-678.

Ha, I. D., Jeong, J.-H. and Lee, Y. (2017). Statistical modelling of survival data with random effects: h-likelihood approach. Springer.

Accelerating Clinical Development by Incorporating Historical Controls in Proof of Concept Studies

Ivan SF Chan, AbbVie Inc.; **Zailong Wang**, AbbVie, Inc.; **Lanju Zhang**, AbbVie, Inc.

In an effort to increase the efficiency of drug development, innovative clinical trial designs have been proposed where historical control data are used in the design and analysis of a new trial. Including historical controls can potentially reduce the number of subjects needed in the new trial and increase the precision of estimating treatment effects. However, there is potential bias that may lead to false decision making. In this talk, we will first discuss the opportunities and challenges in using historical controls in modern clinical trials. Then we will propose a practical approach to leveraging historical controls in the design and analysis of a new trial. In particular, we will discuss several design considerations including compatibility of historical controls, methodology for analysis, evaluation of potential bias and impact on power and sample size. Finally, we will illustrate this proposed approach with an example in an immunology study.

Invited Session	1:30-3:00PM, August 29, 2019	Room 554, 5F

IS17: Statistical Methodology for the Comparative Assessment of Quality Attributes

Can Statistical Inference Improve the (Bio-) Similarity Exercise?

Kit C.B. Roes, Radboud University Medical Centre

The recent (draft) EMA reflection paper aims to establish a framework and a common language to define the appropriate role of inferential statistical methods for comparative evaluation of drug product's quality characteristics. This is a scientific challenge in the sense of answering the different comparison questions based on relevant data and using appropriate (statistical) methods, and understanding the (all) sources of variability involved. It is also a challenge in bridging understanding of all problems at hand and limitations involved between regulatory statisticians, quality assessors and their industrial counterparts. The common interest is that improved assessment of similarity at the quality level will benefit patients, as well as drug development efficiency.

A contentious point in the multidisciplinary discussions on this topic, is the extent to which (formal) statistical inference can add value to the similarity exercise. Formal statistical inference is typically associated with hypothesis testing, p-values and equivalence testing in the setting of similarity. In this presentation, I hope to clarify the broader interpretation of inference that necessarily is (also) applicable to similarity exercises on quality characteristics of drug products. From that, some potential approaches are sketched, with reference to the ongoing revision of the reflection paper.

Similarity Assessment of Quality Attributes: The Calculation of Operating Characteristics to Compare Different Statistical Approaches

Thomas Stangler, Novartis Pharma AG; **Martin Schiestl**, Sandoz

The comparison of quality attributes is a key element in the evaluation of biosimilars and manufacturing process changes for biological medicines. Different statistical approaches were proposed to facilitate such evaluations, however, there is no regulatory consensus on a quantitative and scientifically justified definition and underlying hypothesis of a statistically equivalent quality and it's implication on the operating characteristics of different approaches. This talk proposes a hypothesis for statistically equivalent quality which is in line with current regulations. It also describes a tool which allows comparisons of different statistical approaches or tests by calculating the operating characteristics for false accept and false reject rates of a claim for statistically equivalent quality. These error rates should be as low as possible to allow a meaningful application of a statistical approach in regulatory decision making. The described tool is therefore suitable to compare different statistical approaches for their suitability and may also facilitate the discussion and development of statistical approaches for comparing quality attributes in similarity assessments in general.

Analytical Similarity and Comparability: What is the Question?

Bruno E. Boulanger, PharmaLex Belgium; **Timothy T. Mutsvari**, PharmaLex Belgium

For comparability or analytical biosimilarity purposes, it is the new process and its future capability that should be evaluated, i.e., the risk of producing batches outside defendable limits is key in the evaluation. It is not the new drug product that is evaluated, but rather the performance of the process making the product.

The proposed objective is to consider the overall capability of new processes and the metric proposed to measure the distance between old and new processes is the patient risk or the probability to produced batches of drug product outside the acceptance limits.

To evaluate the capability of a new process it is required to have specifications or acceptance limits, existing or to be defined. For comparison pre-post change, the previous specifications or the previous control limits could be used for evaluating the current and future capability of the new process. For biosimilar products, such pre-existing specifications or control limits based on a long history of production don't exist and have be built and defended using a limited amount of material.

With this objective in mind, Bayesian thinking and statistics is a natural and easy-to-implement solution. With Bayesian statistics it is natural and easy to derive the predictive distribution given the data, meaning that it is possible to compute the predictive probability that new batches will be outside acceptance limits or specifications.

In the presentation, we'll provide an overview of a global Bayesian strategy and evaluate the relative operating characteristics and advantages or disadvantages.

Topic-Contributed Sessions

Topic-contributed Session	10:45AM -12:15PM, August 27, 2019	Room 554, 5F

TC16: The Implementation of ICH-E17in Asian Regions

Taiwan CDE's Experience to Review MRCT Results

I-Chun Lai, Taiwan Center for Drug Evaluation

In Taiwan, the two key steps of new drug review are bridging study evaluation (BSE) and new drug application (NDA). When doing BSE and NDA review, we follow ICH E17 Guideline. The main focus of BSE is to understand whether there is clinically relevant difference between East Asia/East Asian and non East Asia/non East Asian. Population or region can be pooled based on similar intrinsic and/or extrinsic factors. The main focus of NDA review is benefit risk evaluation. If the benefit risk ratio is reasonable to the overall trial population, and the results of overall trial population can be extrapolated to Taiwanese, then the drug can be approved in Taiwan.

MRCT has become the majority of clinical trials conducted in Taiwan in recent years. In 2004, about 52% of clinical trials conducted in Taiwan were MRCTs, the percentage has markedly increased and reached to 70% in 2017. Taiwan CDE has a lot of experience in reviewing MRCTs and also participated in ICH E17 Expert Working Group since it was established in 2014. After ICH E17 Guideline reached step 4 in 2017, Taiwan CDE continued participation in the ICH E17 Implementation Working Group.

In this presentation, examples of reviewing MRCTs will be discussed. With the implementation of ICH E17 guideline, we hope that ethnic factors could be prospectively considered during early drug development stage and more Taiwanese could be enrolled both at the exploratory and confirmatory trials.

Key Principles of the ICH E17 and Their Implementation

William W. Wang, Merck Sharp & Dohme Corp.

The ICH guidance entitled "E17 General Principles for Planning and Design of Multi-Regional Clinical Trials (MRCTs)" was finalized in Nov 2017. This talk will focus on the implementation of ICH E17, with focus on the 7 key principles of effective MRCT designs. We will further discuss the linkage of these 7 principles with the 7 habits of highly effective people (Stephen Covey 1989) and how statisticians can play leadership roles in the ICH E17 implementation.

Topic-contributed Session	10:45AM -12:15PM, August 27, 2019	Room 555, 5F

TC02: New Developments for Statistical Methods in Personalized Medicine

Recommendation for Thirty-Day Rehospitalization Reduction

Menggang Yu, University of Wisconsin – Madison

Thirty-day rehospitalization rate is a well-studied and important measure reflecting the overall performance of health systems. A multitude of efforts have been initiated to reduce avoidable rehospitalizations. These transitional care (TC) programs typically ask nurses to follow-up with patients after the hospitalization to manage issues and reduce the risk of rehospitalizations during health care transitions. As rehospitalization is a complex process that depends on many factors, it is unlikely that these interventions are effective for all patients across a diverse population. In this paper, we consider individualized intervention or treatment recommendation rules (ITRs) aimed at maximizing overall treatment effectiveness. We investigate our approach in a setting where patients are divided into two diagnosis related groups, medically complicated and uncomplicated. As the treatment effects can greatly vary between the two groups, we allow our recommendation rules to be potentially different for different groups. In particular, our approach can take care of scale differences in treatment effects and utilize a tuning parameter to drive the similarity of the estimated ITRs between groups. Computation is realized by morphing our problem into solved forms and a wrapper R package is developed for our proposed treatment recommendation framework. We conducted extensive evaluation through both simulation studies and analysis of a TC program.

Comparative Intervention Scoring for Assessing Heterogeneity of Long-term Health System Intervention Effects

Jared Huling, Ohio State University; **Menggang Yu**, University of Wisconsin-Madison; **Maureen Smith**, University of Wisconsin-Madison

With the growing cost of health care in the United States, the need to improve efficiency and efficacy of the delivery of care has become increasingly urgent. To this end, there have been widespread efforts to design and implement interventions which coordinate the typically fragmented care of complex patients, yet the effectiveness of such interventions in practice has been mixed. A common thread among successful care coordination interventions is the targeting of patients likely to benefit for enrollment, however, there is little guidance toward effectively doing so. In this work we seek to fill this gap by introducing a procedure to estimate personalized scores which characterize differential benefit of long-term health system interventions. As patients tend to respond differently over time, our approach allows the differential effects of an intervention to vary with time and encourage these effects to be more similar for closer time points. We utilize our approach to construct personalized enrollment decision rules for a complex case management intervention in a large health system and demonstrate that the enrollment decision rules result in improvement in health outcomes and care costs.

Change-Point Detection for Infinite Horizon Dynamic Treatment Regimes

Yair Goldberg, Technion; **Moshe Pollak**, The Hebrew University of Jerusalem; **Alexis Mitelpunkt**, Tel-Aviv University andTel-Aviv Sourasky Medical Center; **Mila Orlovsky**, Tel-Aviv Sourasky Medical Center; **Ahuva Weiss-Meilik**, Tel-Aviv Sourasky Medical Center; **Malka Gorfine**, Tel-Aviv University

A dynamic treatment regime is a set of decision rules for how to treat a patient at multiple time points. At each time point, a treatment decision is made depending on the patient's medical history up to that point. We consider the infinite-horizon setting in which the number of decision points is very large. Specifically, we consider long trajectories of patients' measurements recorded over time. At each time point, the decision whether to intervene or not is conditional on whether or not there was a change in the patient's trajectory. We present change-point detection tools and show how to use them in defining dynamic treatment regimes. We demonstrate the utility of the proposed change-point detection for detecting sepsis in preterm infants in the intensive care unit and detection of a change in glucose levels of a diabetic patient.

Multi-Category Individualized Treatment Regime Using Outcome Weighted Learning

Xinyang Huang, East China Normal University; **Yair Goldberg**, Technion - Israel Institute of Technology; **Jin Xu**, East China Normal University

Individualized treatment regimes (ITRs) aim to recommend treatments based on patient-specific characteristics in order to maximize the expected clinical outcome. Outcome weighted learning approaches have been proposed for this optimization problem with primary focus on the binary treatment case. Many require assumptions of the outcome value or the randomization mechanism. In this paper, we propose a general framework for multi-category ITRs using generic surrogate risk. The proposed method accommodates the situations when the outcome takes negative value and/or when the propensity score is unknown. Theoretical results about Fisher consistency, excess risk and risk consistency are established. In practice, we recommend using differentiable convex loss for computational optimization. We demonstrate the superiority of the proposed method under multinomial deviance risk to some existing methods by simulation and application on data from a clinical trial.

Topic-contributed Session	1:30 -3:00PM, August 27, 2019	Room 554, 5F

TC07: Multiplicity Issues in Complex Clinical Trials

An Enhanced Mixture Method for Constructing Gatekeeping Procedures in Clinical Trials

Thomas Brechenmacher, IQVIA Japan; **Alex Dmitrienko**, Mediana Inc.; **George Kordzakhia**, U.S. Food and Drug Administration; **Eiji Ishida**, U.S. Food and Drug Administration

It is increasingly common to encounter complex multiplicity problems with several multiplicity components in confirmatory Phase III clinical trials. These components are often based on several endpoints (primary and secondary endpoints) and several dose-control comparisons. When constructing a multiplicity adjustment in these settings, it is important to control the Type I error rate over all multiplicity components. An important class of multiple testing procedures, known as gatekeeping procedures, was derived using the mixture method that enables clinical trial sponsors to set up efficient multiplicity adjustments that account for clinically relevant logical relationships among the hypotheses of interest. An enhanced version of this mixture method is introduced to construct more powerful gatekeeping procedures for a broad class of multiplicity problems that are very common in Phase III clinical trials. Examples are provided to illustrate the new method and to compare the performance of gatekeeping procedures derived using this method and other available methods.

Group Sequential Designs for Clinical Trials with Multiple Survival Endpoints

Kentaro Sakamaki, Yokohama City University

Group sequential designs are used in some oncology clinical trials with multiple primary endpoints, such as progression free survival (PFS) and overall survival (OS). It is simple to use an alpha sending function for each primary endpoint based on an alpha level split by the Bonferroni procedure in order to control the type I error rate in group sequential designs for clinical trials with multiple survival endpoints. However, there remain challenges on analysis times and alpha spending functions for multiple survival endpoints. Analysis times can be different between multiple survival endpoints because information fractions are based on numbers of events. If analysis times are determined by one primary endpoint where events occur frequently, an adaptive change of analyses for another primary endpoint may be operationally useful. For example, we can consider that OS is analyzed at interim and final analysis using an updated alpha level when statistical significance is achieved for PFS, but OS is analyzed only at final analysis using the originally assigned alpha level when statistical significance is not achieved for PFS. However, it is known that an adaptive change occurs an inflation of the type I error rate. In this talk, we propose group sequential designs for clinical trials with multiple primary survival endpoints and evaluate the type I error rate and power of proposed approaches through simulation studies.

Practical Strategies for Testing Co-primary Endpoints in Group-sequential Clinical Trials

Koko Asakura, National Cerebral and Cardiovascular Center; **Toshimitsu Hamasaki**, National Cerebral and Cardiovascular Center; **Frank Bretz**, Novartis Pharma AG

In this presentation, we discuss the methods for testing hypotheses associated with two co-primary endpoints in group-sequential clinical trials comparing the test intervention to the control intervention. Co-primary means that a trial is designed to evaluate a joint effect on all of the primary endpoints, and failure to demonstrate statistical significance on any single endpoint implies that test intervention's effect to control cannot be concluded. We consider a specific situation where the null hypothesis for one endpoint has been rejected at some interim look and the null for the other endpoint have not yet rejected, but the result has been promising. Our question is how the null for the other endpoint can be tested after then. We discuss design adaptations for testing the null for the other endpoint, with: (1) increasing the number of analyses, (2) decreasing the number of analyses and (3) changing the alpha allocation to subsequent analyses. We investigate the operating characteristics of these adaptations in term of the power, Type I error, and sample sizes. We illustrate the methods using a real clinical trial.

A Case Study of Refining Testing Strategy Using Graphical Approach

Naoko Kataoka, Novartis Japan K.K.; **Jiawei Wei**, Novartis; **Eva Hua**, Novartis

A clinical trial usually includes multiple endpoints to be tested, and we need to consider how to test the several hypotheses considering multiplicity. As the endpoints are correlated each other, we should consider the overall power (to reject all key hypotheses under multiplicity consideration) of the study.

The situation is more complex when we have multiple doses and comparators. A graphical approach is available for developing and evaluating hierarchal multiple analysis strategies, this approach is most helpful when the analysis plan is complex due to splitting of the overall alpha among several endpoints. Simple modifications of the initial graph can easily create different variations of a test strategy, therefore help comparison among the variations. A case study will be presented how we refined our testing strategy in a confirmatory trial using graphical approach, based on comparing the overall power of the possible testing strategies.

Topic-contributed Session	1:30 -3:00PM, August 27, 2019	Room 555, 5F

TC01: Biostatisticians Role in Innovative Trial Design in the New Era of Drug Development

Improving the Assessment of Probability of Success in Late Stage Drug Development

Lisa Hampson, Novartis; **Steffen Ballerstedt**, Novartis; **Björn Bornkamp**, Novartis; **Björn Holzhauer**, Novartis; **Joseph Kahn**, Novartis; **Markus Lange**, Novartis; **Wen-Lin Luo**, Novartis

There are several steps to confirming the safety and efficacy of a new medicine. A sequence of trials, each with its own objective, is usually required. Quantitative risk metrics can be useful for informing decisions about whether a medicine should transition from one stage of development to the next. Traditionally, pharmaceutical companies have used cross-industry success rates to estimate the probability of obtaining regulatory approval. Project teams then typically apply subjective adjustments to reflect project-specific information. However, this approach lacks transparency and fails to make full use of data from previous clinical trials. We describe a quantitative Bayesian approach for calculating the probability of success (PoS) at the end of Phase II, which incorporates internal clinical data, cross-industry success rates, and expert opinion or external data if needed. Using an example, we illustrate how PoS can be calculated accounting for differences between our Phase II data and future Phase III trials, and how the sensitivity of PoS to assumptions can be evaluated and communicated. We will also discuss how PoS can be used in earlier stages of development to inform decision making.

A Statistical Framework for Quantitative Decision Making in Early Clinical Development

Weidong Zhang, Pfizer Inc.

Drug development is becoming increasingly costly due to the high attrition rate. To reduce cost and improve success rate, evidence based decision making is critical for clinical development. Key decisions in early drug development include but not limited to 1) do we see enough evidence to confirm proof of mechanism (PoM) of the drug 2) do current efficacy and safety support further development? 3) can we modify (stop, accelerate or enroll more patients) an ongoing trial with accumulating data? Traditional decision making rarely depends on formal quantification and relies mostly on an ambiguous process involving subjective expert opinion. Therefore, it is critical to develop a quantitative and evidence-based decision making framework which can provide a full picture of benefit and risk.

In this presentation, we developed a dual-criterion approach that can not only be used to objectively evaluate pharmacology for PoM declaration, but also can be extended to other clinical studies with efficacy as the objective. In addition, we adopted a Bayesian frame work that took into account all information including data collected from current study and data from previous studies. Metrics to measure drug modulation effect, efficacy and statistical confidence will also be discussed. Objective assessment can be made using this approach and false decisions can be significantly minimized.

Comparison of Frameworks for Tipping Point Analyses

Achim Guettner, Novartis**; Summer Xia**, Novartis

The final concept paper “Addendum to Statistical Principles for Clinical Trials on Choosing Appropriate Estimands and Defining Sensitivity Analyses in Clinical Trials” emphasizes the importance for having a framework on sensitivity analyses. Sensitivity analyses can cover the presentation of results of analyses in which different choices and departures of assumptions are considered. These analyses include assumptions made on missing data across the different endpoints. This can be addressed with tipping point analyses. Different frameworks for tipping point analyses will be presented and discussed. Some approaches allow the inclusion of covariates in the analyses. However, this inclusion adds on additional complexity to the analysis and interpretation of the results. In addition to the statistical properties, the ease of clinical interpretation also needs to be considered in the choice of the methodology. Simulations to compare the impact of covariates on the tipping point analysis were conducted and results will be presented. Further considerations of tipping point analyses relevant at study planning will become obvious.

Dynamic Bayesian Decision Making in Early Phase Trials Using Historical Information

Fan Xia, BeiGene Ltd.; **Xiao Lin**, BeiGene Ltd

Considering the uncertainty of the treatment effect and observed data variation due to a limited sample size, Go/No-Go decision making is a widely-concerned issue in early phase trials of exploratory nature. Traditionally, many frameworks for early phase decision making have been proposed based on frequentist methods. However, in the rapid and competitive drug development, decision making more need to involve accumulated information in a dynamic fashion, and is implemented not only at design stage but also during the conduct of the trial and at the analysis stage for exploratory trials. A Bayesian method is more feasible to update decision model with refreshed internal and external data in the way. In this presentation, we proposed a Bayesian meta-analytical-predictive (MAP) method to derive informative priors from historical data, which can be combined with new data collected from current trials at the time of decision making. Decision criteria then can be defined based on posterior distributions. Considering prior-data conflict, which is unavoidable despite careful selection of historical studies, we proposed a robust version of MAP method to discount historical information based on the degree of confidence in the relevant of the historical data. At last, to facilitate this method, we have developed an user-friendly R shiny App to help the implementation of such methodology.

Topic-contributed Session	3:30 -5:00PM, August 27, 2019	Room 554, 5F

TC11: Advances in Design and Analysis of Clinical Studies that Incorporate Internal and External Data Sources

Target Population Statistical Inference with Data Integration across Multiple Sources: An Approach to Mitigate Information Shortage in Rare Disease Clinical Trials

Yang Song, Vertex Pharmaceuticals Inc.; **Xihao Li**, Harvard University

For clinical trials in rare disease areas, a major challenge is the limited amount of available information for making robust statistical inference due to low disease prevalence. External data sources present data integration opportunities to enhance statistical inference. We propose an intuitive integrated inference framework to integrate information from all relevant data sources and make inference on the treatment effect over a specific target population. The method is easily implemented and extended with modern machine learning tools. It is complemented by a variance estimation procedure to facilitate statistical inference. The proposed method is shown to have good statistical properties with both theoretical development and simulation studies. We argue that the integrated inference framework not only provides an intuitive and coherent perspective for a wide range of clinical trial inference problems but also has broad application areas in clinical trial settings and beyond, as a quantitative data integration tool for making robust inference in a target population precise manner for policy and decision makers.

Estimation of the Effect of the NSAID Celecoxib on the Risk of Cancer using Electronic Healthcare Record Data

Tasuku Okui, Kyushu University Hospital; **Naoki Nakashima**, Kyushu University Hospital; **Atsushi Takada**, Kyushu University Hospital; **Fumi Takahashi-Yanaga**, University of Occupational and Environmental Health; **Toshiyuki Sasaguri**, Kyushu University

Celecoxib is a nonsteroidal anti-inflammatory drug (NSAID), and from its pharmacological mechanism of action, its possible effect in preventing the risk of cancer development is expected. Although several pharmacoepidemiologic studies have been conducted to establish the safety of celecoxib, no study of its effect on the risk of cancer has been conducted yet. We conducted this study to verify the effects of celecoxib on the risk of development of cancer using real- world data: electronic healthcare record (EHR) data. EHR data are useful for extracting the detailed medical history of each patient.

The research subjects were patients who were prescribed celecoxib or the control drug, Loxoprofen, at Kyushu university hospital. We extracted several characteristics and the medical history of the patients like the age, date of consultation, prescribed drugs, and history of illness from the EHR database system. As the characteristics of the patients differed between the two drug groups and the number of variables were many, we constructed a predictive model for the groups and estimated the propensity score for each patient using a machine learning model. Then, we conducted survival analysis for the occurrence of cancer using a cox proportional hazards model with inverse probability of treatment weighting and g-estimation using a structural accelerated failure time model.

The results revealed a significantly increased incidence rate of cancer, especially in women, in the celecoxib group. This result has not been indicated by other studies, and we propose to conduct a validation study using data from other university hospitals.

Biomarker-integrated Clinical Trials with Threshold Selection and Enrichment

Xiaofei Wang, Duke University School of Medicine; **Ting Wang**, University of North Carolina; **Stephen L. George**, Duke University School of Medicine

A biomarker with potential to guide treatments is often measured on a continuous scale and optimal cutoffs for patient subgrouping are not available at the time of trial design. Biomarker-driven RCT designs have been proposed to simultaneously search for optimal cutoffs for subgroup identification and to evaluate treatment effects therein. The existing designs often fail to achieve the study objectives, as the threshold searching algorithm is prone to large variability and bias and the patients who benefit from the experimental therapy are incorrectly identified. We study a novel design, in which the study hypotheses may be refined

based on cumulative trial data and adaptive algorithms are used to search for tentative ‘optimal’ cutoffs. Enrichment strategies are then used to ensure optimal efficiency in testing the refined hypotheses and sufficient accuracy in identifying the final ‘optimal’ cutoffs at the end of the trial. We study the design for continuous, binary, and time-to-event endpoint in a multiple-stage framework. Issues related to trial implementation and monitoring, conditions for efficiency gain, and properties of treatment effect estimators under the new design will be discussed.

Topic-contributed Session	3:30 -5:00PM, August 27, 2019	Room 555, 5F

TC09: Adverse Events in Clinical Trials and Post-Marketing Pharmacovigilance

Rationale and First Results from the SAVVY Project

Regina Stegherr, University of Ulm

The analysis of adverse events (AEs) is an essential part of the assessment of safety in the evaluation of new therapies in clinical trials. There is general agreement among stakeholders including regulators, payers, industry, healthcare professionals and patients that improvements in the evaluation of a drug's safety would benefit all. Although statistical methodologies have advanced, are implemented in standard software (such as R and SAS) and therefore readily available, statistical AE analyses are often too simplistic and therefore potentially misleading. The ``Survival analysis for AdVerse events with VarYing follow-up times'' (SAVVY) project aims to improve the analyses of adverse event data in clinical trials through the use of survival techniques appropriately dealing with censoring, competing risks (events before AE occurrence) and varying follow-up times. In an empirical meta-analytic study including randomized controlled clinical trials by a number of sponsor companies, the SAVVY project investigates the potential impact that improved AE analyses might have on the conclusion of the safety assessment compared to the current standard of incidence proportions or incidence densities.

In this talk the rationale, statistical concept and first results of the SAVVY project will be presented.

A Bayesian Meta-analytic Approach for Safety Signal Detection in Randomized Clinical Trials

Motoi Odani, Ono Pharmaceutical Co., Ltd.; **Satoru Fukimbara**, Ono Pharmaceutical Co., Ltd.; **Tosiya Sato**, Kyoto University School of Public Health

Meta-analyses are frequently performed on adverse event data and are primarily used for improving statistical power to detect safety signals. However, nave pooling of adverse event data across multiple clinical trials is still commonly used. We sought to propose a Bayesian hierarchical meta-analytic approach based on consideration of a hierarchical structure of reported individual adverse event data across multiple randomized clinical trials. To develop our meta-analysis model, we extended an existing three-stage Bayesian hierarchical model by including an additional stage of the clinical trial level in the hierarchical model; this generated a four-stage Bayesian hierarchical model. We applied the proposed Bayesian meta-analysis models to published adverse event data and to a simulation study motivated by the case example to evaluate the characteristics of three alternative models. Comparison of the results from the Bayesian meta-analysis model with those from Fisher’s exact test after nave pooling showed that more individual adverse events under the certain body system were detected with regard to association with treatment in the Bayesian meta-analysis model than in Fisher’s exact test. Based on the simulation results, the Bayesian meta-analysis model could balance the false detection rate and power to a better extent than Fisher’s exact test. Our proposed meta-analysis models considered trial effects to avoid confounding by trial and borrowed strength from both within and across body systems to obtain reasonable and stable estimates of an effect measure by considering a hierarchical structure of adverse events.

Modified Bayesian Confidence Propagation Neural Network for Signal Detection Analysis

Keisuke Tada, Sanofi. K.K.; **Kazushi Maruo**, University of Tsukuba; **Naoki Isogawa**, Pfizer R&D Japan; **Yusuke Yamaguchi**, Astellas Pharma Inc.; **Masahiko Gosho**, University of Tsukuba

On the development of a new drug, evaluations for safety are usually insufficient due to the limited number of treated patients in clinical trials. After the drug launched, information regarding safety of the drug should be kept collecting. An approach to detect potential adverse drug reaction (ADR) in spontaneous reporting system databases is called “signal detection”. A Bayesian Confidence Propagation Neural Network (BCPNN) is one of the signal detection methods used in World Health Organization Uppsala Monitoring Center. We modified the BCPNN to increase the sensitivity for detecting potential ADR. In BCPNN, information component (IC) is defined as an index of disproportionality between the observed and expected number of reports mentioning a drug and an event. A positive IC indicates stronger association between the drug and the event than expected. Our proposed method is to adjust the IC value by borrowing information of events occurred in drugs defined as similar to the target drug. We compared the performance of our

method with that of the traditional BCPNN through a simulation study. In the simulation, sensitivity of the proposed method is much higher than that of the traditional BCPNN method in case that the differences of IC between a target drug and similar drugs are from 0 to 2.

Topic-contributed Session	10:45AM - 12:15PM, August 28, 2019	Room 554, 5F

TC17: Innovative and Strategic Thinking in Pediatric and Early Drug Development based on Baysian Hierarchical Model

Leveraging Available Information in Pediatric Trial Designs and Analyses using Bayesian Modeling

Susan Q. Wang, Boehringer-Ingelheim Pharmaceuticals, Inc.

There are many challenges facing the development of a pediatric drug. Innovative and unconventional statistical methods are imperative in assisting drug development in this area. Bayesian statistical methods have been used more and more often to integrate historical information via using a prior. The prior can be built applying Bayesian hierarchical model, power prior, or a meta-analytic predictive prior. In certain situations, the empirical Bayes approach may also be considered. In this presentation, we discuss the application of the Bayesian meta-analytic approach to integrate information from multiple existing trials into a new pediatric trial. We used simulations to quantify the gains.

A Robust Bayesian Approach for Using Co-data in Phase I Oncology Trials: An Application to Bridging Studies

Haiyan Zheng, Newcastle University; **Lisa Hampson**, Novartis Pharma AG, **Thomas Jaki**, Lancaster University

We propose a robust Bayesian hierarchical model to leverage co-data, e.g., the animal and human toxicity data collected from heterogeneous sources, in phase I bridging trials. Embedded with translation factors, parameters of the dose-toxicity models in any tested animal species and patient subgroups can be interpreted on a common scale, say, the equivalent human dosing scale. Prior distributions are specified to capture uncertainty about the magnitude of the translation factor appropriate for each species. The study-specific model parameters are assumed to be exchangeable within an animal species, and a “supra-species” random-effects distribution is stipulated to model the species-specific population parameters for increased borrowing of information across species. To relate the information from animals and humans, we assume the dose-toxicity parameters of a phase I oncology trial are drawn from bivariate normal distributions, of which the means are consistent with those describing toxicity in animals species studied so far. More specifically, for dose-toxicity parameters that underpin each patient subgroup, the probability of consistency is split to reflect the prior scepticism about their similarity with those in a specific animal species, while the possibility of inconsistency is also enabled for robust inferences. The proposed methodology is illustrated using data examples and simulations. Numerical results suggest that our methodology improves the precision of estimates of the toxicity rates when the co-data are consistent, and it discounts inconsistent co-data quickly avoiding excessive shrinkage of parameters towards the population average estimate for an extreme patient subgroup.

Introduction to the Bayesian Early-Phase Seamless Transformation (BEST) Platform Design

Jiaying Lyu, Fudan University; **Wentian Guo**, Laiya Consulting Inc.; **Yuan Ji**, Laiya Consulting Inc. and the University of Chicago

Recently the seamless strategy of early phase drug development has led to successful development of many breakthrough cancer immune therapies such as Pembrolizumab. With the release of the draft guidance of multiple expansion cohorts in first-in-human trials by FDA in August 2018, it is desirable to consider seamless phase 1a/1b trials where expansion of multiple indication cohorts follows a phase 1a dose-escalation trial seamlessly. This type of seamless strategy is more complex than the traditional designs and requires advanced statistical modeling and designs. In this presentation, we introduce a novel Bayesian early-phase seamless transformation (BEST) platform design that combine a phase 1a dose-escalation stage, phase 1b cohort expansion stage with multiple indication cohorts, or even a phase 2 stage. The BEST platform utilizes a Bayesian hierarchical model that can improve the overall study power in terms of selecting the promising dose and indication for later-stage drug development. We demonstrate the benefits of the BEST design using simulation studies based on an ongoing clinical trial.

A Case study: A Bayesian Type Adaptive Dose Finding in a Clinical Trial with Multiple Agents

Wenxiao Zhou, Beigene, Ltd.

Combination therapy with more than one agent nowadays has become a trend in cancer treatment and research and there is a growing interest in identifying the maximum tolerated dose combination (MTDC) based on multiple agents, especially when the agents with potentially different biological characteristics may exhibit enhanced clinical treatment effect when used together.

A real-world case is recently raised where three different agents, of which one agent's safety profile is unclear and the other two being relatively well-known, are being considered in order to find both optimal single-agent dose and optimal dual-agent dose combinations in one single Phase I clinical trial.

With this regard, we propose a two-stage, Bayesian type, adaptive dose-finding design where an initial dose escalation stage (i.e. a BLRM design with late onset AE controlled in the mono search and a 3+3-like quick combo exploration along the diagonal of dose combination matrix) is followed by an adaptive and systematic combo-finding stage where a Bayesian logistic regression model is used and ultimately suggests dual-agent combinations that are considered MTDC.

Under this design, it can a) incorporate a line of on-going single-agent trial, b) insert unplanned dose combination(s) at potentially the end of trial, c) potentially initiate cohorts in parallel to save time and resources, and d) recommend one combo for each level of agent 2. Our method is compared to other previously proposed methods in a similar setting in terms of accuracy, safety, and trial speed.

Topic-contributed Session	10:45AM - 12:15PM, August 28, 2019	Room 555, 5F

TC08: Innovative Methods to Support the Development of New Pediatric Medicines

A Clinician's View of the Importance of Pediatric Extrapolation

Robert M. Nelson, Johnson & Johnson

Achieving the vision of a world in which children have access to safe and effective medicines for immune mediated diseases at the same time as their parents requires the innovative use of extrapolation to reduce the need for generating pediatric evidence. Extrapolation is an inductive inference that extends known experience and/or data into an area not known or previously experienced to arrive at a credible, but inherently uncertain knowledge of the unknown area. This presentation explores the scientific and clinical assumptions that behind using prior knowledge to draw conclusions about the efficacy of drugs for use in children. Extrapolation also involves a value judgment about the tolerable uncertainty (i.e., the type 1 error) appropriate for a specific pediatric development program. The tolerable uncertainty concerning the decision to approve a new drug (i.e., efficacy) often is framed by the risk of being wrong across an entire population. The tolerable uncertainty concerning the decision to expose a child to the risks of a new drug in a research protocol is framed by the risks and potential benefits to that individual child. Arguably, the lack of safe and effective treatment options for a smaller pediatric population with severe life-threatening disease argues for a level of tolerable uncertainty closer to that of individual clinical decisions. The clinical and ethical issues concerning the use of extrapolation in pediatric drug development will be illustrated conceptually through the application of Bayesian statistical methods and the use of prior knowledge in model-based exposure-response analysis.

Extrapolating Information from Adult to Paediatric Studies: a Comparison of Methods

Juan Jose Abellan-Andres, GlaxoSmithKline; **Fi Guillard**, GlaxoSmithKline; **Christina Filmore** GlaxoSmithKline; **Nirav Ratia**, GlaxoSmithKline; **Min Sun**, GlaxoSmithKline; **Ohad Amit**, GlaxoSmithKline; **Nicky Best**, GlaxoSmithKline

The development of new medicines for paediatric population is confronted with ethical and practical concerns that may even threat the feasibility of clinical studies in children. Common diseases in adults may be rare in children; well-established clinical procedures in adults may not be ethical or even possible in paediatric subjects, etc. However, when the paediatric development of a new drug starts, typically there is already valuable information from the development of the same drug in adults, which can be used to support the limited evidence that can be obtained in children. The usefulness of extrapolation is reflected in guidance documents from regulatory agencies (e.g. the EMA and the US FDA), which also include some discussion on general statistical principles for extrapolation. A number of approaches have been suggested in the literature that can be used to extrapolate information from adults to support the assessment of the efficacy of a new drug in paediatric. But we are still in the process to understand how these methods perform in practice. In this presentation we will show the results of a simulation study to compare some of those approaches using well-known operating characteristics under several scenarios including both superiority and non-inferiority settings. The approaches we investigated include the robust mixture prior, the commensurate prior and a few methods from the power-prior family. Our results suggest the robust mixture prior and the commensurate prior are the most appealing overall, with the former being slightly preferred because of greater transparency.

Current Situation of Pediatric Drug Development and Evolving Discussion on Extrapolation in Japan

Hidefumi Nakamura, National Center for Child Health and Development; **Yasuhiko Imai**, Japan Pharmaceutical Manufactured Association

There is no regulation to mandate pediatric drug development in Japan. The effective incentive for pediatric drug development is believed to be the premium to promote development of new drugs and to eliminate off-label use in National Reimbursement System. This premium is applied in relation to the decision by the Expert Panel on unapproved drugs and off-label drugs with high medical need. Another incentive is an extension of reexamination period.

PMDA has established a Pediatric Working Group focusing on Pediatric issues in drug development and approval in 2011. The group identifies the obstacles for pediatric development and discuss with stakeholders to resolve problems and accelerate pediatric development. PMDA is also attempting to create and utilize models to identify pediatric dosage efficiently. The Committee on Pharmaceutical Affairs of the Japan Pediatric Society (JPS) is also actively involved in drug development issues in children. Starting October 2017, the JPS Drug Development Network (JPeDNet) was established supported by a grant from the Japan Agency for Medical Research and Development. The JPeDNet functions closely with the Pediatric Clinical Trial Network that consists of 41 hospitals including most major children's hospitals in Japan.

There is a lag in pediatric drug development between Japan and the EU/US. The discussion on methodological aspects of extrapolation is not only on extrapolation from adults to children, but also on extrapolation of real-world evidence in actual clinical setting. Discussion on utilization of extrapolation between academia and industries has been recently started. Update on current discussion will be introduced.

The Challenge of Implementing Bayesian Methods in Paediatric Studies: Our Experience at UCB

Rosalind J. Walley, UCB Pharma; **Foteini Strimenopoulou**, UCB Pharma

Bayesian methods are increasing used in the pharmaceutical industry. Using an informative prior for placebo is commonplace in proof of concept studies. Bayesian methodology has been developed for the paediatric setting. And yet to-date there appear to be very few completed regulatory Bayesian paediatric studies. This is intriguing when we consider that for some rare diseases, single arm trials are acceptable, and in the paediatric setting, "full extrapolation" (i.e. not carrying out a paediatric efficacy study) is sometimes acceptable. Both of these scenarios might be viewed as Bayesian positions, with extremely strong priors on placebo and the paediatric treatment effect respectively.

In this talk we present some of the challenges we have faced at UCB when implementing these methods using examples from epilepsy and immunology. We will focus on how Bayesian methodology fits with the concept of extrapolation as described in regulatory guidances, how historical data has been selected to construct informative priors when there are different sources of available information, the difficulty of estimating study-to-study variation and how the impact of using informative priors has been assessed.

Topic-contributed Session	10:45AM - 12:15PM, August 28, 2019	Room 555, 5F

TC14: Other Way forward for Design, Summary Measures, and Estimands in Survival Clinical Trials

Hazards of Proportional Hazards Assumption and Small Number of Events: Actual Clinical Problems in Oncology and Alternative Ideas

Shogo Nomura, National Cancer Center

Traditionally, randomized controlled trials (RCTs) for cancer patients have used survival endpoints (e.g., progression-free survival and overall survival) as a primary endpoint and have applied standard survival analysis methods such as the log-rank test and the Cox regression model. The performance of these methods largely depends on the proportional hazards assumption, and, to guarantee the asymptotic normality of test statistics, accumulation of events above a certain level is a must. When these assumptions are violated, the log-rank test is no longer most powerful, estimated hazard ratio (HR) is not a useful and interpretable summary measure of treatment effect. In recent oncology RCTs, we have increasingly faced with these situations, for example: (i) Kaplan-Meier curves crossed or separated after a time-lag when trial objective is to show a superiority of actively-developed oncology new agents; (ii) primary analysis had to be done using a data with far smaller number of events than researchers expected at the planning phase.

In this presentation, trial designs and analysis methods under non-proportional hazards or small number of events are overviewed focusing on recent activities (e.g., some research letters for medical top journals [e.g., *the New England Journal of Medicine*] and a consensus made by FDA-initiated collaborative working group about non-proportionality of hazards). Especially, the pros/cons of two actively-discussed methods (MaxCombo test and restricted mean survival time) will be discussed by motivating actual superiority and non-inferiority trials. My presentation ends with an introduction of talks in this session.

Using Restricted Mean Survival Time in a Non-Inferiority Trial

Isao Yokota, Hokkaido University; **Yukari Uemura**, The University of Tokyo Hospital

Non-inferiority (NI) trials are intended to confirm that the new therapeutic arm is no worse than a standard arm, but has less intensity of treatment, less toxicity, and improve quality of life. On NI trials with time-to-event outcome targeted on low-risk patients, relatively lower event rate leads to huge sample size or a wider NI margin of hazard ratio (HR). Furthermore, the NI margin of HR should be interpreted in accordance with the baseline hazard function. As an alternative approach, the NI margin of restricted mean survival time (RMST) has been discussed, which may also be described by an absolute difference.

In this presentation, we will discuss the translation and statistical efficiency of RMST and HR margin. If Weibull distribution for the survival function is assumed, RMST can be calculated explicitly. We will show the correspondence of NI margin of HR, RMST and the survival rate that are calculated from the specific survival function under several scenarios with varied shape and scale parameters. Through the simulation study, the power and empirical relative efficiency of the NI margin of HR and RMST will also be shown.

As a result, interpreting HR as the ratio of median survival time, there was a broad gap between the difference of median survival time and RMST. The efficiency of RMST is as well as or better than HR in proportion to the width of the NI margin.

Pairwise Pseudolikelihood Estimation of an Average Hazard Ratio under Nonproportional Hazards

Tomohiro Shinozaki, Tokyo University of Science

Hazard ratio is a popular summary measure that is commonly reported in survival clinical trials thanks to its identifiability and attainable efficiency with flexible models under proportional hazards. Under nonproportional hazards, however, a single hazard ratio estimate will lack the well-defined estimand; many of recent applied research, including treatments with delayed effect or heterogeneous effects on distinct subgroups, have also thrown doubt on its interpretability. Moreover, nonproportional hazards indicate the presence of different “susceptibility” for treatment among patients, which induces selection bias in hazard itself and deprive it of causal interpretation. In this talk, we show that under nonproportional hazards, it is possible to “average” time-varying hazard ratios into a single well-defined summary measure that is

causally interpretable as a “concordance” probability of paired data, known as a probabilistic index. The probabilistic index interpretation of a time-constant hazard ratio has been well described, while its counterpart of the “average” time-varying hazard ratios is less well-known. We describe the existent, heuristic estimator using weighted Cox models and the more straightforward, proposed estimator using pairwise-stratified weighted Cox models for possible pairs in the sample. The proposed estimator would be justified by pairwise pseudolikelihood theory. We will provide numerical results in simulation and real clinical trial data, both of which lack proportionality of hazards. Even if the proportional hazards assumption is violated, hazard ratios will remain to be one of the useful summary measures of treatment effects insofar as adequate “average” is taken over time.

A Design Consideration on RCTs Assessing Superiority of Immuno-oncology Agents: Sample Size Determination and Monitoring

Takahiro Hasegawa, Shionogi & Co., Ltd.

Immuno-oncology agents often require time to elicit an immune response and a delayed treatment effect is expected. Because of this delayed effect, the agent does not initially influence the survival curves of the trial results. If the treatment is effective, the survival curves will separate once the agent’s effect has established in superiority studies. The delayed treatment effect violates the assumptions of the Cox proportional hazards model and substantially reduces the statistical power of conventional methods such as the standard log-rank test. Therefore, the use of a weighted log-rank test with the Fleming-Harrington class of weights has been proposed as an alternative analysis method for survival endpoints.

First, we give a flexible survival model in immuno-oncology studies that considers not only the delayed treatment effect but also the long-term survivors which were shown in some actual studies. Moreover, we present a method for calculating the sample size for the weighted log-rank test with the Fleming-Harrington class of weights under assumption of the proposed flexible survival model. The impact of delayed effect timing on both the choice of the Fleming-Harrington’s weights and the increment in the required number of events is discussed. In addition, it should be noted that an information fraction for the weighted log-rank test is not proportional to the number of events. Therefore, we propose the corresponding semiparametric information fraction, which can be used for group sequential designs based on the error spending function approach and is easily calculated without unblinding treatment assignment.

Topic-contributed Session	1:30 - 3:00PM, August 28, 2019	Room 554, 5F

TC05: Machine Learning Methods for Improving Clinical Decision-Making and Precision Health Care

Learning Optimal Individualized Treatment Strategies from Randomized Trials and Electronic Health Records

Yuanjia Wang, Columbia University; **Peng Wu**, Columbia University; **Donglin Zeng**, University of North Carolina at Chapel Hill; **Haoda Fu**, Eli Lilly

Individualized treatment rules or recommendations (ITRs) tailor medical treatments according to patient-specific characteristics in order to optimize patient's outcome. Data from randomized controlled trials (RCTs) are used to infer valid ITRs using statistical and machine learning methods. However, RCTs are usually conducted under specific inclusion/exclusion criteria, thus limiting the generalizability of ITRs to a broader real-world patient population. On the other hand, since patient's electronic health records (EHRs) document treatment prescriptions in the real world, transferring information in EHRs to RCTs, if done appropriately, could potentially improve the performance of ITRs, in terms of precision and generalizability. In this work, we propose a new domain adaptation method to learn ITRs by incorporating evidence from EHRs. Due to the presence of unmeasured confounding in EHRs, we do not directly learn the optimal ITR from the combined EHR and RCT data. Instead, we first pre-train "super" features from EHRs that summarize physicians' treatment decisions and patient's observed benefits in the real world, which are likely to be informative of the optimal ITRs. We present theoretical justifications and conduct simulation studies to demonstrate the performance of our proposed method. Finally, we apply our method to transfer information learned from EHRs of type 2 diabetes (T2D) patients to improve learning individualized insulin therapies from an RCT.

Application of Machine Learning to Real World Data

Shintaro Hiro, Pfizer R&D Japan; **James J. Kim**, Pfizer Inc.

Real world data (RWD) and real world evidence (RWE) are now playing an increasingly important role in healthcare decisions. RWD is any healthcare data for decision making that is not collected in conventional randomized controlled trials. It includes electronic health records, claims or billing activities, and product or disease registries. In an epidemiological study using RWD, the first step often is to define the target disease or outcome of interest based on terms from the master codes of diseases and/or medical procedures. ICD-10 (International Statistical Classification of Diseases and Related Health Problems, 10th revision) is one such set of master codes. In actual clinical settings, however, the choice of disease name varies across different physicians and hospitals. In addition, disease histories differ among patients. Therefore, it is extremely challenging to make exhaustive definition of the target disease or outcome. We propose to find a common set of medical records for the target disease/outcome using machine learning approach, namely, the gradient boosting method. Availability of such common set of records for various diseases/outcomes of interest would be helpful for researchers embarking on studies using RWD.

Joint Variable Screening in Accelerated Failure Time Models

Jinfeng Xu, University of Hong Kong; **Yixin Fang**, AbbVie

Variable screening has gained increasing popularity in high-dimensional survival analysis. Most existing methods for variable screening with survival data suffer from that variable importance is assessed based on marginal models that relate the time-to-event outcome to each variable separately, implying that the relevance of one variable is examined when other variables are excluded. These methods will preclude variables that only manifest their influence jointly and may retain irrelevant variables that are correlated with relevant ones. To circumvent these difficulties, we propose a new approach to evaluating joint variable importance in censored accelerated failure time models. We establish the sure screening properties of the proposed approach and demonstrate its effectiveness through simulation studies and a real data application. A novel stability selection-based procedure is also proposed for tuning.

Topic-contributed Session	1:30 - 3:00PM, August 28, 2019	Room 555, 5F

TC04: Machine Learning Methods for Improving Clinical Decision-Making and Precision Health Care

Bayesian Dose- Finding Phase I Trial Design Incorporating Historical Data from a Preceding Trial

Kentaro Takeda, Astellas Global Development Inc.; **Satoshi Morita**, Kyoto University

We consider the problem of incorporating historical data from a preceding trial to design and conduct a subsequent dose - finding trial in a possibly different population of patients. In oncology, for example, after a phase I dose-finding trial is completed in Caucasian patients, investigators often conduct a further phase I trial to determine the maximum tolerated dose in Asian patients. This may be due to concerns about possible differences in treatment tolerability between populations. In this study, we propose to adaptively incorporate historical data into prior distributions assumed in a new dose-finding trial. Our proposed approach aims to appropriately borrow strength from a previous trial to improve the maximum tolerated dose determination in another patient population. We define a “historical-to-current (H-C)” parameter representing the degree of borrowing based on a retrospective analysis of previous trial data. In simulation studies, we examine the operating characteristics of the proposed method in comparison with 3 alternative approaches and assess how the H-C parameter functions across a variety of realistic settings.

Selection of Robust Meta-Analytic-Predictive Priors based on the Evaluation of Operating Characteristics for Proof-of-Concept Studies

Yi Cheng, China Novartis Institutes for Biomedical Research Co., Ltd

Historical data are always utilized for designing clinical trials. Especially for phase II proof-of-concept studies, the use of historical data for determining “Go/No Go” decision based on Bayesian methods is widely accepted. However, despite a careful selection of the historical trials, the new trial may have some unsuspected features which make it inconsistent with the historical trials. Schmidli H, Gsteiger S, Roychoudhury S, et al (2014) proposed a method to construct robust meta-analytic-predictive prior, which includes a weighted vague conjugate prior and can alleviate prior-data conflicts. In this presentation, we will first review some applications of this method in proof-of-concept studies. Then, we will focus on the choice of the weights for the vague conjugate prior, and will evaluate the design operating characteristics with various weights. At last, some recommendations will be provided based on the simulation results.

Utility of Bayesian Single-Arm Design in New Drug Application for Rare Cancers in Japan

Akihiro Hirakawa, The University of Tokyo

Investigational drugs for rare cancers are often approved based solely on a single-arm phase II trial that primarily evaluates response rate in Japan. Such trials typically use a fixed sample size determined on the basis of the frequentist manner. However, since predicting the speed of patient enrollment is challenging because of the disease rarity, the time needed to complete the enrollment of the fixed number of patients is prolonged in some cases. A Bayesian design without fixing the sample size is useful for single-arm phase II trials of rare cancers. However, the arbitrariness of prior distribution specifications and the frequentist operating characteristics are regulatory issues. In this presentation, we share the experience of conducting a Bayesian single-arm phase II trial of an investigational drug in patients with sarcoma for new drug application, along with the discussion with regulatory agency in Japan.

Use of Co-data for Interim Analysis in Clinical Trials

Tomoyuki Kakizume, Novartis Pharma K.K.; **Heinz Schmidli**, Novartis Pharma AG; **Beat Neuenschwander**, Novartis Pharma AG; **Sebastian Weber**, Novartis Pharma AG

Clinical trials often include interim analyses, at which time it is decided whether to stop or continue the trial. To make a good decision, all relevant data should be used. These consist not only of the interim data from the ongoing trial, but also of co-data. Co-data comprises all relevant (historical and concurrent) data external to the trial. The robust meta-analytic-predictive (MAP) approach can be used to synthesize the

co-data and robustly predict parameters of the ongoing trial, within a Bayesian framework. For implementation of the robust MAP approach in an interim analysis of a clinical trial, the R package RBesT (https://CRAN.R-project.org/package=RBesT) may be used. A phase III trial with interim analysis for futility stopping will be used for illustration, where the co-data consist of two historical trials and interim data from an additional ongoing phase III trial.

Topic-contributed Session	3:30 - 5:00PM, August 28, 2019	Room 554, 5F

TC13: Data Science for Medicine

Integrative Analysis of Genetic, Transcriptomic and Functional Data in Identification of Potential Driver Genes in Tumors

Chen Suo, Fudan University; **Wenjiang Deng**, Karolinska Institute; **Yudi Pawitan**, Karolinska Institute

Many studies have identified aberrations related to the pathogenesis and prognosis, broadly classifying neuroblastoma patients into high- and low-risk groups, but predicting tumor progression and clinical management of high-risk patients remains a big challenge.

We integrate gene-level expression, array-based comparative genomic hybridization and functional gene-interaction network of 145 neuroblastoma patients to detect potential driver genes. The drivers are summarized into a driver-gene score (DGscore) for each patient, and we then validate its clinical relevance in terms of association with patient survival. Focusing on a subset of 48 clinically defined high-risk patients, we identify 193 recurrent regions of copy number alterations (CNAs), resulting in 274 altered genes whose copy-number gain or loss have parallel impact on the gene expression. Using a network enrichment analysis, we detect four common driver genes, ERCC6, HECTD2, KIAA1279, EMX2, and 66 patient-specific driver genes. Patients with high DGscore, thus carrying more copy-number-altered genes with correspondingly up- or down-regulated expression and functional implications, have worse survival than those with low DGscore ($P = 0.006$). Furthermore, Cox proportional-hazards regression analysis shows that, adjusted for age, tumor stage and MYCN amplification, DGscore is the only significant prognostic factor for high-risk neuroblastoma patients ($P = 0.008$).

Integration of genomic copy number alteration, expression and functional interaction-network data reveals clinically relevant and prognostic putative driver genes in high-risk neuroblastoma patients. The identified putative drivers are potential drug targets for individualized therapy. The integrative pipeline has also been generalized and applied to breast cancers and lung cancers.

A Bivariate Shared Parameter Model for Intensive Longitudinal Data Subject to Informative Missing

Xiaolei Lin, Fudan University; **Xiaolei Xun**, Fudan University; **Robin Mermelstein**, University of Illinois at Chicago; **Donald Hedeker**, **University of Chicago**

In this paper, we address the problem of informative missing in the context of intensive longitudinal studies, where each study unit gets measured intensively over time and intermittent missing is usually present. We present a bivariate shared parameter model approach that links the primary bivariate longitudinal outcomes with potentially informative missingness by a common set of random effects that summarizes subjects' specific traits in terms of their mean (location) and variability (scale). The primary outcome vector, conditional on the random effects, are allowed to be correlated, exhibit heterogeneity with respect to both the mean and within subject variance, and are further assumed to be independent of the missing process. Unlike the previous methods which largely rely on numerical integration or approximation, we estimate the model by a full Bayesian approach using MCMC. An adolescent mood study example is illustrated together with a series of simulation studies. Results in comparison to the conventional but possibly inappropriate approaches suggest that accounting for informative missingness by sharing both mean and variance model parameters, as well as joint modeling the bivariate outcome vector can significantly improve the model fit yet provide the benefit of understanding how missingness can affect the inference for the primary outcome and how the bivariate outcomes are correlated over time.

Data Processing and Data Analysis with Real World Big Data

Kazuo Ishii, Kurume University

In recent years, along with the development of information and communication technologies (ICT), global data traffic is growing significantly. Most of these big data has been disposed without using. It is promised that revolutionary innovation is brought about by effective use of Big Data. In this study, I will focus on Real World Big Data (RWD) in clinical sites, such as medical prescription and Diagnosis Procedure Combination (DPC) data. We analyzed the DPC-based commercial RWD, MDV medical database, consists of 197,645

subjects' data, to investigate side effects by anti-cancer drugs. Then, we established the method to remove (or reduce) the effect of confounding factors and to efficiently detect risk factors of side effects by big data processing. We could show some risk factors, such as sex, age, BMI and some comorbidities, were related to side effects of anti-cancer drugs. Here, I will describe what is Real World Big Data, and its processing and analysis methods with our practical examples.

Sequential Adaptive Subject and Variable Selection for Generalized Estimating Equation Methods

Zimu Chen, University of Science and Technology of China; **Zhanfeng Wang**, University of Science and Technology of China**; Yuan-Chin Chang**, Institute of Statistical Science, Academia Sinica

We present an aggressive sampling strategy, using the ideas of active learning methods in machine learning literature, to accelerate the estimating methods with generalized estimating equations (GEE), which is one of the popular statistical methods for analyzing correlated or highly stratified multiple-response data, and can handle many types of unmeasured dependence between outcomes. We integrate the adaptive sampling and variable selection features into a sequential procedure for modeling the correlated or highly stratified multiple-response responses data. Besides reporting the statistical properties of the proposed procedure, we also use both synthesized data and real data sets to demonstrate the usefulness of our method.

Topic-contributed Session	3:30 - 5:00PM, August 28, 2019	Room 555, 5F

TC10: Designing Clinical Trials with Recurrent Events

Designing a Trial in Multiple Sclerosis with Relapses as Endpoint

Isao Tsumiyama, Novartis Pharma K.K.

Patients with relapsing multiple sclerosis (RMS) experience neurologic symptoms, so-called “relapses” in a recurrent way. One of the treatment goals in RMS is to reduce the frequency and severity of relapse. In clinical trials for RMS, relapse is evaluated commonly as the primary endpoint. Frequencies of relapses are summarized as event rates, usually called annualized relapse rates (ARR). In this talk, it is introduced how and what to consider to evaluate relapses in RMS clinical trials.

Introducing a Clinical Trial Sample Size Calculation: Experiences in Hemophilia A using recurrent events

Wataru Ohtsuka, Chugai Pharmaceutical Co., Ltd.;

[Background]HAVEN1 is a PhaseIII trial (funded by F.Hoffmann-La Roche and Chugai) of emicizumab, a bispecific antibody that bridges activated factor IX and factor X to restore the function of missing activated factor VIII that is needed for effective hemostasis in persons with hemophilia A. Bleeding is a recurrent event and characterized by an over-dispersion compared to a Poisson distribution often considered when modelling count data. Sample size calculation under an over-dispersed assumption is therefore required.

[Methods]The sample size calculation was based on the number of bleeds over time (bleed rate) with emicizumab prophylaxis(treatment, λt) versus no-prophylaxis(control, λc) following a negative binomial(NB) distribution with shape parameters(γt and γc).

[Results]Considering enrollment feasibility, 45 patients under an allocation ratio of 2:1 (emicizumab prophylaxis:no-prophylaxis) achieved >95% power under λt/λc is 4/18 using East, Version 6. It was assumed that the patients were followed up to 0.5 year.

However standard tools do not account for different follow-up time among patients. Power was also estimated by simulation on the basis of NB regression model including an offset variable to account for variable follow-up times; the sample size had >95% power at the 2-sided 0.05 level of significance.* At the primary analysis, the annualized bleed rates were 2.9(emicizumab prophylaxis) versus 23.3(no-prophylaxis), representing a statistically significant reduction(P<0.001). Observed over-dispersion was in line with the model assumptions(Oldenburg J, et al. N Engl J Med 2017;377:809-18).

[Conclusion]The model achieved convergence for all the bleed endpoints. It was successfully used in a clinical trial for recurrent events under an over-dispersed assumption.

Power and Sample Size Calculation in Clinical Trials with Over-dispersed Count Data

Masataka Igeta, Hyogo College of Medicine; **Shigeyuki Matsui**, Nagoya University Graduate School of Medicine

In comparative clinical trials with recurrent events, such as exacerbations of chronic obstructive pulmonary disease (COPD), over-dispersion needs to be appropriately incorporated in both design and analysis to control the type I error rate and power. Plausible miss-specifications of the variance function (derived from quasi-Poisson or negative binomial models) may be handled at the data analysis stage by using a robust test to control the type I error rate. However, it is difficult to control the power under variance miss-specifications because of limited prior knowledge on the over-dispersion at the design stage. In this paper, we introduce a procedure of sample size determination incorporating variance miss-specifications for stable power control. An application to a randomized clinical trial to confirm the suppressive effect on COPD exacerbations is provided. An extension to repeated count data would also be discussed.

Adaptive Designs for Clinical Trials with Recurrent Events

Tim Friede, University Medical Center Göttingen

Recurrent events are common endpoints in clinical trials, including hospitalizations in heart failure, exacerbations in chronic obstructive pulmonary disease, and relapses in multiple sclerosis. These endpoints are over-dispersed, i.e. their variance is larger than their mean. In this presentation, we start by considering blinded nuisance parameter sample size reestimation (e.g. Friede and Schmidli, 2010; Schneider et al, 2013) and blinded continuous monitoring of the information (Friede et al, 2019), which work with non-comparative data. Furthermore, we will summarize some recent findings on group-sequential designs for studies with overdispersed recurrent events (Mtze et al, 2018a,b). This is joint work with a number of colleagues including Heinz Schmidli and Tobias Mtze (both Novartis, Basel, Switzerland).

1. Friede T, Hring DA, Schmidli H (2019) Blinded continuous monitoring in clinical trials with recurrent event endpoints. Pharmaceutical Statistics 18: 5464.
2. Friede T, Schmidli H (2010) Blinded sample size reestimation with count data: Methods and applications in multiple sclerosis. Statistics in Medicine 29: 1145-1156.
3. Mtze T, Glimm E, Schmidli H, Friede T (2018) Group sequential designs for negative binomial outcomes. Statistical Methods in Medical Research (in press).
4. Mtze T, Glimm E, Schmidli H, Friede T (2018) Group sequential designs with robust semiparametric recurrent event models. Statistical Methods in Medical Research (in press).
5. Schneider S, Schmidli H, Friede T (2013) Blinded sample size reestimation for recurrent event data with time trends. Statistics in Medicine 32: 54485457.

Topic-contributed Session	10:45AM - 12:15PM, August 29, 2019	Room 555, 5F

TC03: Regulatory Submissions in Electronic Format

Study Data Technical Rejection Criteria, Validation, and Self-Check Worksheet

Ethan Chen, US Food and Drug Administration

Study Data Standards listed in the FDA Data Standards Catalog are required for clinical and nonclinical studies that started after December 17, 2016 (for ANDA, NDA and BLA) or December 17, 2017 (for Commercial IND). Through the technical rejection process, FDA can reject an application because of its technical deficiencies, based on the severity of the eCTD validation criteria. FDA conducted an analysis on submissions that contain study data that were already received by the Agency to assess conformance rates to Technical Rejection Criteria for Study Data (TRC). Submissions received between December 18, 2016 to March 31, 2018 (December 18, 2017 to March 31, 2018 for Commercial IND submissions) showed that about 32.0% of all submissions with study data were received with critical errors (i.e. submissions with 1734 and/or 1736 errors).

Based on findings from this analysis, FDA updated TRC to provide more clarification and developed supporting tools to help Industry meet study data requirements. FDA also conducted an analysis on submissions that contain study data received in 2018 to generate additional findings to the baseline analysis. These efforts are expected to improve conformance rates over time by making it clearer and easier for Industry to meet FDA's study data requirements.

Differences between FDA and PMDA for E-data Submission

Masato Suzuki, MSD K.K.

Japanese pharmaceuticals will face a steep cliff in 2020. Why? We will not be allowed to file any new products to (Pharmaceuticals and Devices Agency) PMDA without e-data. In some cases, the e-data submitted to the Food and Drug Administration (FDA) cannot be reused for the PMDA submission without an update because of regulatory differences between PMDA and FDA.

NMPA Reform and Keytruda Filing in China

Jing Zhang, Merck Sharp & Dohme Corp.

Expanding on recent reforms allowing innovative drugs to be approved on the basis of overseas clinical trial data, China's National Medical Products Administration (NMPA) (formerly the China Food and Drug Administration or CFDA) has created special channels for the approval of new pharmaceuticals subject to "urgent" clinical needs. In late April 2018, NMPA utilized its new Conditional Approval mechanism to grant marketing authorization to Gardasil 9 human papilloma virus (HPV) vaccine after a mere nine days of review, with the consideration of overseas clinical trial data.

Under this exciting condition, we need to think about how to prepare our eSub packages efficiently and how to use the global data effectively. If global data did not include China, CDE (Center for Drug Evaluation) may consider the conditional approval and sponsor needs to provide two commitment studies, one as primary proposal and the other one acts as back up. In complicated situations, it may depend on other study results and depends on results of futility analysis Pooled Asian subgroup analysis may be needed.

How to present safety data is also important for CDE. They prefer to separate monotherapy and combination. They also prefer to integrate China safety into global database instead of separate safety summary, except there are different safety profiles for China and global.

Information Requests during an FDA Review

Hong Qi, Merck Sharp & Dohme Corp.; **Lei Xu**, Merck & Co., Inc.; **Mary N. Varughese**, Merck Sharp & Dohme Corp.

Filing a marketing application is a pivotal and exciting milestone for the long-term effort of product development. Even though there are industry standards to follow, submissions vary among products, indications, and sponsoring companies. Regardless of the extraordinary efforts in preparing the submission package, it is common for regulatory agencies to issue information requests (IRs) from the pre-supplemental biologics license application (sBLA) meeting, during the application review, and labeling process. IRs may arise from different aspects including the target indication, patient population, the drug safety profile, the reviewers' scientific interest on getting further information on the potential benefits of the medicine, additional case-report forms, and even the collected previous therapies not included in the ADaM datasets.

This paper will describe the data preparation and submission pertaining to IRs received from the pre-sBLA, during the sBLA review and the labeling, with the focus on the approaches we utilize, and some thoughts on future strategies.

Topic-contributed Session	10:45AM - 12:15PM, August 29, 2019	Room 509, 5F

TC15: Real-Word Data: Implications and Challenges for Medical Product Development

Validity and Reliability of Real-World Data for Medical Product Development

Taro Shibata, National Cancer Center

For rare diseases and other cases in which it is difficult to conduct traditional and/or orthodox randomized controlled clinical trials, efforts have been made in recent years to explore the use of data from existing patient registries for marketing approval applications of medical products.

For each patient registry, the purpose and status of its design/operation determine whether or not its data can be used for approval applications. In addition, when using patient registry data for approval applications, the reliability of the data needs to be ensured at a level considered sufficient in light of its purpose. Further, it is not necessarily adequate to use the same level of reliability as that employed in clinical trials; instead, the level should be specific to patient registries.

Against this background, the concept of data reliability assurance in the use of patient registries for approval applications was discussed by an AMED research project team. Then, focusing on the use of patient registry data in a new drug application dossier, the project team presented a proposal regarding the requirements for the design and operation of patient registries that should be met when these registries are used for approval applications of medical products.

This presentation will provide a summary of the proposal documents.

Statistical Methods in Use of Real-world Data in Medical Product Development: Propensity-Based Methods

Hisateru Tachimori, National Center of Neurology and Psychiatry

Propensity-based methods are increasingly applied to causal inference with observational data. The propensity score allows us to design and analyze an observational (nonrandomized) study so that it mimics some of the particular characteristics of a randomized controlled trial (Austin, P., 2011). The propensity-based methods are one of the most promising methods to make evidences from real-world data in medical product development, though we should make a consensus in using the methods in medical product development. In this presentation, we would like to share the points to consider to use the propensity score methodology properly in medical product development and discuss the prospective design of observational studies with propensity score methodology.

Instrumental Variable Analysis in Clinical Trials Incorporating Patient Registry Databases as Controls

Yukari Uemura, The University of Tokyo Hospital; **Tomohiro Shinozaki**, Tokyo University of Science; **Noriko Tanaka**, Tokyo Metropolitan Geriatric Medical Center

While well-controlled and well-conducted randomized clinical trials are viewed as a gold standard in the safety and effectiveness evaluation of premarket medical products, there has been an increasing interest in exploring opportunities to utilize data collected in a high-quality patient registry databases to form the whole or a part of the control group, especially for rare disease. Among numerous kinds of biases that may occur using a registry database as a control group, one of the major concerns is the presence of potential bias due to non-comparability between the treatment group and the control group, that is, confounding bias. To address confounding bias in observational studies, outcome regression or propensity score methods has been widely applied. However, these conventional approaches will fail to estimate causal effects when large or unknown sources of unmeasured confounding are suspected (as is often the case of nonrandomized studies). In this talk, we will review an alternative analysis technique, instrumental variable (IV) method, which can estimate the effect in the presence of unmeasured confounding. However, the validity of IV analysis requires the specific conditions and assumptions that are not testable from observed data; hence, there is a tradeoff between conventional and IV analyses regarding the assumptions that make estimates valid in clinical trials using patient registry databases. We will specifically focus on the points to

consider in applying IV method for marketing approval applications of pharmaceuticals, to the combined data of a clinical trial (treatment group) and a registry database (control group).

Topic-contributed Session	1:30 - 3:00PM, August 29, 2019	Room 555, 5F

TC12: Opportunities and Challenges for the Use of Parametric Longitudinal Modelling in Drug Development

Opportunities and Pitfalls in the use of Nonlinear Mixed-Effects Models for Leveraging Longitudinal Information in Drug Development

Andrew C. Hooker, Uppsala University

A number of recent studies have shown that the use of nonlinear mixed-effects models to analyze longitudinal data can dramatically increase the power of clinical studies in drug development trials compared to standard statistical analyses at end of treatment. These pharmacometric model approaches attempt to characterize the longitudinal dose-exposure-response (DER) relationships, which can be useful for identifying drug effects, selection of efficacious doses, as well as understanding and characterizing other aspects of a pharmaceutical compound, for example, drug-drug interactions. Further, the models can be used in planning and optimizing new trials, and can be an integral part of the decision-making process in drug development and usage. However, one problem with pharmacometric analyses is the potential bias introduced through model building and the use of the “wrong” model in the model-based approaches. By properly accounting for model and parameter uncertainty (for example with the use of model averaging techniques and/or adaptive optimal design) one can alleviate these issues in the evaluation of clinical trial data, the design of future trials and in decision making.

Quantifying and Addressing Model Uncertainty on Longitudinal Data in the Design and Analysis of Clinical Studies

Tobias Mielke, Janssen Pharmaceuticals; **Vladimir Dragalin**, Janssen Pharmaceuticals

Adaptive study designs are considered as one source for increased efficiency in drug development. A common complication in the implementation of adaptive designs is the limited information available at the timing of interim analyses. There is not much room for increased operational efficiency in situations with fast patient recruitment and long time to endpoint. As a result, all available data need to be used optimally for effective decision making at an early time point. Although longitudinal data are typically collected in clinical studies, primary analyses are frequently based on model-free landmark analyses or MMRM approaches. Parametric longitudinal modelling could increase the information available for decision making. However, parametric modelling comes at the risk of potential bias introduced from the underlying modelling assumptions. Considerations on potential model-misspecification should hence be included into the design process for studies utilizing longitudinal modelling. Concerns on model uncertainty have been widely discussed for dose-finding studies in the scientific literature. The methodology can be generalized to longitudinal modelling, covering model selection approaches, model-based contrast tests and/or (Bayesian) model averaging approaches.

The effects of model uncertainty on interim decision making will be discussed in this presentation. Different parametric and semi-parametric longitudinal modelling approaches will be evaluated for this purpose. The presented approaches will be compared in their ability of mitigating concerns on the true underlying longitudinal model, while increasing efficiency in decision making.

Leveraging Parametric Longitudinal Modeling to Improve Drug Development Efficiency

Olga V. Marchenko, Bayer; **Jose C. Pinheiro**, Janssen

Although data in drug development are often collected longitudinally, frequently primary and key secondary analyses rely on a single post-randomization time point (e.g., change from baseline, response at Week X, etc.). While focus on a single time point may be driven by regulatory requirements in a “confirmatory” setting, longitudinal modeling utilizing the totality of the observed data offers the opportunity to achieve substantial efficiencies in data analysis compared to current prevailing practices at the “learn” stage of drug development. Those efficiencies may translate into faster and more accurate decision making (e.g., interim analysis adaptation and futility rules, adaptive dose selection), increased statistical power and/or estimation precision, and combinations of these, ultimately leading to increased probability of program success (technical, regulatory and commercial), while reducing development costs. This presentation will discuss

and illustrate some of the opportunities associated with the broader use of parametric longitudinal modeling, but also the caveats associated with it, such as the need for additional assumptions about the data.

Calibrated Predictions of Survival based on Tumor Size Dynamics and New Lesions in Lung Cancer via Joint Modeling Approach

Katsuomi Ichikawa, AstraZeneca K.K.; **Jacqueline Buros-Novik**, Icahn School of Medicine at Mount Sinai; **Mario Nagase**, AstraZeneca; **Diansong Zhou**, AstraZeneca; **James Dunyak**, AstraZeneca; **Nidal Al-Huniti**, AstraZeneca

BACKGROUND: Tumor burden and growth rate correlate strongly with progression-free survival (PFS) and overall survival (OS) and are the criteria for RECIST-based surrogate endpoints in oncology. A statistical framework to predict at progression in Non-Small Cell Lung Cancer (NSCLC) patients is needed.

METHODS: This work develops a joint model of disease progression and OS that incorporates longitudinal tumor burden, appearance of new lesions, and changes in therapy following progression in NSCLC patients to interrogate the components of RECIST and to predict PFS. We considered a nonlinear exponential growth model for time-dependent tumor growth and the patient-treatment-course-specific parameters are informed by a subset of the patient-treatment-course-specific covariates. We estimate the baseline hazard using a cubic spline with 6 knots, and numerically integrate the hazard to estimate patient-specific survival using Gauss-Kronrad quadrature with 13 quadrature points. The model is estimated using Stan.

RESULTS: The model was fitted to data from 434 NSCLC patients treated with gefitinib or carboplatin+paclitaxel ('IPASS', NCT00322452), and further evaluated on 102 EGFR+ NSCLC patients treated with gefitinib ('IFUM', NCT01203917). The model generated predictions in alignment with observed data and recapitulates RECIST outcomes with nearly 90% accuracy. The sum-of-longest-diameters (SLD) and new lesions were, as expected, associated with OS (hazard ratios: 1.96 and 6.40). Surprisingly, the rate of change in SLD showed little association with OS (hazard ratio 1.00).

CONCLUSION: This joint model generates well calibrated predictions of PFS and OS. It offers insights to the relative predictive value of -components of RECIST in NSCLC.

Topic-contributed Session	1:30 - 3:00PM, August 29, 2019	Room 509, 5F

TC06: Utilization of Subgroup and Casual Inference towards Personalized Medicine

Identifying Targeted Patients Population in Major Depressive Disorder by Enhanced Enrichment Design

Peter Zhang, Otsuka Development & Commercialization Inc.

Despite advances in clinical trial design, failure rates near 80% in phase 2 and 50% in phase 3 have recently been reported. The challenges to successful drug development are particularly acute in central nervous system trials such as for pain, schizophrenia, mania, and depression because high - placebo response rates lessen assay sensitivity, diminish estimated treatment effect sizes, and thereby decrease statistical power. This paper addresses the importance of rigorous patient selection in major depressive disorder trials through an enhanced enrichment paradigm. This approach led to a redefinition of an ongoing, blinded phase 3 trial algorithm for patient inclusion (1) to eliminate further randomization of transient placebo responders and (2) to exclude previously randomized transient responders from the primary analysis of the double blind phase of the trial. It is illustrated for a case study for the comparison between brexpiprazole + antidepressant therapy and placebo + antidepressant therapy. Analysis of the primary endpoint showed that efficacy of brexpiprazole versus placebo could not be established statistically if the original algorithm for identification of placebo responders was used, but the enhanced enrichment approach did statistically demonstrate efficacy. Additionally, the enhanced enrichment approach identified a target population with a clinically meaningful treatment effect. Through its successful identification of a target population, the innovative enhanced enrichment approach enabled the demonstration of a positive treatment effect in a very challenging area of depression research.

Double-robust inference for differences in restricted mean lifetimes using pseudo-observations

Sangbum Choi, Korea University; **Taehwa Choi**, Korea University

When comparing survival times between two treatment groups, restricted mean lifetime, defined as the expectation of the survival function restricted to a prespecified time point, is often of direct interest as it is easily understood by clinician investigators and does not require restrictive assumptions, such as proportionality. If the treatments are not randomized as in observational studies, covariate adjustment is needed to account for treatment imbalances in confounding factors. In this article, we propose a simple pseudo-value approach to estimate the difference of the restricted mean lifetime between two groups while accounting for confounders, which can be used as a metric for average causal effect (ACE). The proposed method combines two general approaches, (i) group-specific regression models (Q-model) for the time-to-event and covariate information, and (ii) inverse probability of treatment assignment weights (A-model), where the restricted mean lifetimes are replaced by the corresponding pseudo-observations for unknown survival quantities.

We show that the proposed estimator is double robust in the sense that it is consistent if at least one of the two working models remains correct. Simulation studies are conducted to assess its finite-sample performance and the method is applied to the GSE dataset.

Fortified Robust Estimate of Rx Effects in Nonrandomized External Control Trial, Subgroup and RWD Analysis

Ming Tan, Georgetown University

Causal inference has become a research focus with increasing availability of real world data (RWD) in drug development and clinical research. RWD has begun to influence trial design and augment randomized controlled trial (RCT). The method is motivated by the design of an ongoing trial to compare relapse-free survival (RFS) rate at 3 years between locally treated high-risk ocular melanoma patients on adjuvant combination immunotherapy versus a matched contemporaneous control because RCT is not possible. The external control is obtained from RWD with the same enrollment criteria. Then an accurate estimate of the treatment effect is important. Given the enrollment criteria for treatment and control are the same, we propose a causal inference method that not only has the doubly robustness property but also procedure but also additional robustness by extending the propensity score model and the regression model to

semiparametric models with monotone constraint on the nonparametric parts. Simulation studies are conducted to evaluate the performance of the proposed method and compare some existing methods. Then the method is applied to analyze a real clinical trial data. The method extends to comparing non-randomized groups in subgroup and RWD analysis. This work is in collaboration with A. Yuan, A. Yin, S. Rapisuwon, M. Atkins

Contributed Sessions

CS1	10:45AM - 12:15PM, August 27, 2019	Room 509, 5F

Integration of Elicited Expert Information via a Power Prior in Bayesian Variable Selection: Application to Colon Cancer Data

Sandrine Boulet, INSERM; **Moreno Ursino**, INSERM; **Peter Thall**, M.D. Anderson Cancer Center; **Bruno Landi**, Hpital europen Georges-Pompidou; **Cline Lepre**, Hpital europen Georges-Pompidou; **Simon Pernot**, Hpital europen Georges-Pompidou; **Anita Burgun**, INSERM and Hpital europen Georges-Pompidou; **Julien Taieb**, Hpital europen Georges-Pompidou; **Aziz Zaanan**, Hpital europen Georges-Pompidou; **Sarah Zohar**, INSERM; **Anne-Sophie Jannot**, INSERM and Hpital europen Georges-Pompidou

Context: Building decision support tools in medicine requires identifying relevant variables to model the medical decision. Currently, machine learning techniques are used to select variables used for decision making from patient care data. Along with data-driven analysis, eliciting experts opinion can be useful to model decision making. Thus, combining expert data to observed data into a new variable selection method could identify the whole set of relevant variables for decision making. We propose a method that introduces experts' information into a Bayesian variable selection model, the Stochastic Search Variable Selection (SSVS) model. We consider the context of medical decisions regarding dose adjustment of Irinotecan for the treatment of metastatic colorectal cancer.

Objective: Clinician first provides numerical clinical relevance weights to express their beliefs about the importance of each variable in their decision of dose adaptation. Then, a sample of simulated data is generated from these weights and combined with the observed data via the power prior method. We compare the performance of our method to the SSVS model in our case study.

Method: The elicited weights deeply vary across clinicians. Performance does not depend on the amount of expert information. For the same amount of observed and expert information combined using the power method, we are able to select rare variables with high elicited weight.

Result: We present a Bayesian variable selection method incorporating elicited expert information and observed data. The method selects a set of relevant variables to model the medical decision process.

Assessing Overall Treatment Effect Based on Robust Estimation in Multi-Regional Clinical Trials

Shuhei Kaneko, Novartis Pharma K.K.; **Akihiro Hirakawa**, The University of Tokyo

A use of multi-regional clinical trials (MRCTs) in clinical developments increased to fasten patients' access to new drugs globally. MRCTs inherently assume that treatment effect is consistent across regions; therefore, we generally estimate overall treatment effect by using a weighted mean of regional estimates with proportions of the enrolled patients as weight. However, recent MRCTs occasionally encountered an inconsistency of treatment effect among regions. This study was motivated by the two actual MRCTs showing the inconsistency of treatment effect. In these trials, few regional estimates showed no effect or were harmful despite overall treatment effect was statistically significant, and the weighted mean was heavily influenced by them. To address this, we propose a new method for assessing overall treatment effect estimate that adaptively changes the weights for each region using robust M-estimator. Specifically, we devise the two types of weights for the regional estimates; one is the proportion of the enrolled patients in each region, and another is the difference between the regional and overall treatment effects. The latter weights are controlled by a tuning parameter, and we propose to determine the optimal value of the tuning parameter by minimizing jackknife-based mean square error of M-estimator. We present the proposed estimator provides less biased estimate compared to the weighted mean through simulation studies. We also illustrate the sensitivity assessment of overall treatment effect estimate by applying the proposed estimator to the two actual MRCTs.

Robust Estimates of Regional Treatment Effects in Multiregional Randomized Clinical Trial with Ordinal Response

Chongyang Duan, Southern Medical University; **Ao Yuan**, Georgetown University; **Ming T. Tan**, Georgetown University

Global clinical trials involve multiple regions in the world. Determining the regional treatment effects of a new therapy over an existing one is important for both the sponsors and the regulatory agencies of the regions. Existing methods are mainly for continuous primary endpoint and use subjectively specified models, which may deviate from the true one. Here we consider trials with ordinal response (as primary endpoint). Utilizing recent theoretical advances in ordinal regression models, we develop a robust semiparametric ordinal model for estimating the regional treatment effect, in which the regression coefficients and regional effects are modelled parametrically for ease of interpretation, and the regression link function is specified nonparametrically for robustness. The model parameter is estimated by semiparametric maximum likelihood estimates, and the null hypothesis of no regional effect is tested by the Wald statistic. Simulation studies are conducted to evaluate the performance of the proposed method, compare with the commonly used parametric model, and results show improved performance over the latter. Then the method is applied to analyzing a real multiregional clinical trial with ordinal response.

Sample Size and Power Calculations for Reference-Based Imputation

Kimitoshi Ikeda, AbbVie GK

The handling of missing data in clinical trials is important to appropriately assess the treatment effect in clinical trials. The primary analysis is generally based on the missing at random (MAR) assumption. A MAR analysis provides de jure estimand of the treatment effect which is the effect under the ideal situation that all patients take the assigned study drug throughout the study. A MAR analysis does not provide appropriate estimates in case that patients do not take study drugs after discontinuation due to efficacy or safety reasons.

Reference-Based Imputation (RBI) has been proposed as an approach not assuming MAR. RBI assumes that patients who have discontinued test drug have the same response profile as the response profile of patients in the control group and provides de facto estimand of treatment effect. RBI is attractive because the underlying assumption is interpretable from a clinical point of view and provide conservative estimates of the treatment effect.

In this presentation, we focus on the Copy Reference (CR) and the Jump-to-reference (J2R) of the RBI method and consider the treatment effect and variance estimates. In addition, we show analytic expressions of the treatment effect and derive closed form expressions for power and sample size calculation for CR and J2R approaches. We perform the simulation and examine the performance of power and sample size calculation methods.

Using Cure Rate Models to Characterize Survival Data in Oncology

Kohinoor Dasgupta, Novartis Healthcare Pvt. Lmt.; **Tomas Haas**, Novartis Pharma AG; **Ekkehard Glimm**, Novartis Pharma AG

Cure rate models represent a statistical modelling approach developed to model and predict time-to-event data in situations where it is reasonable to assume there is a subset of patients who will never experience the event of interest and are therefore ‘cured’. The assumption about the presence of ‘cured’ patients is supported when Kaplan-Meier curves show a plateau in their tails, suggesting that the probability of new events becomes close to zero after a certain time point. Cure models can be used to investigate the heterogeneity between patients with cancer who are long-term survivors and those who are not. In adjuvant disease setting, a cure rate model is a reasonable analysis because there is a proportion of patients who will be “cured” by complete resection and will remain free of relapse after surgery. In this presentation, we use cure rate models to describe the effect of a treatment over surgery alone as a proportion of patients will be “cured” by resection. We explore different mixture and non-mixture cure rate models and evaluate the factors influencing survival of cured and non-cured patients with the help of simulated data and clinical trial data from adjuvant melanoma disease. In this context we show that cure rate models characterize the effect of treatment in terms of increasing the cure rate as well as increasing the survival time of the non-cured patients.

Design and Analysis of Biosimilarity with an Estimated Margin on Interval Estimations

Chieh Chiang, National Health Research Institutes; **Chin-Fu Hsiao**, National Health Research Institutes

In the analyst stage for assessing the biosimilarity between an innovative biological product and its biosimilar drug, one criterion might be that the difference of mean values of the two products are compared against an equivalence margin in the form of k times the reference variability. In practice, this margin is unknown and should be estimated. Moreover, the variabilities of the two drugs are different and unknown leads to the leads to the so-called Behrens-Fisher problem — there does not exist an exact test or interval estimation on the mean difference. In this study, we construct two one-sided interval estimations to resolve the issue. Simulation shows the proposed method can control type I error rate and provide a targeted level of power asymptotically.

CS2	8:45- 10:15AM, August 30, 2019	Room 510, 5F

Multiple Imputation with Auxiliary Variables in Longitudinal Clinical Trials: Imputation Models Using Bayesian Lasso and Tree-Based Approaches

Yusuke Yamaguchi, Astellas Pharma Inc.; **Toshihiro Misumi**, Yokohama City University School of Medicine; **Kazushi Maruo**, University of Tsukuba

Multiple imputation is a promising approach for handling missing data, where imputation models are used for multiply imputing the missing values of outcome variables. In longitudinal clinical trials, the imputation model is routinely specified by a linear combination form of the outcome variables (and covariates included in analysis models if necessary). However, in the context of clinical trials, several dozens of variables including post-randomisation data (e.g. treatment compliance and onset of adverse events) are available on hand and these have the potential to be auxiliary variables incorporated into the imputation model. The use of auxiliary variables can be beneficial for making a missing at random assumption more plausible and helping to reduce uncertainty of the imputation, while it is also true that the nave inclusion of too many non-informative auxiliary variables may just add noise rather than being helpful. We address multiple imputation methods using Bayesian lasso and tree-based approaches (random forest and gradient boosting machine), which allow data-driven specification of the imputation model. Specifically, the Bayesian lasso and the tree-based algorithms are built in a procedure of multiple imputation by chained equations. This brings out flexibilities of selecting informative auxiliary variables automatically and improving the predictive accuracy of imputed values. A simulation study suggested the methods achieved more accurate imputation and led to robust estimation of treatment effects with higher power in comparison to conventional approaches. We also found the Bayesian lasso outperformed the others in certain situations. The methods were illustrated through an application to a real example.

A Variable Selection Criterion for Competing Risk Data with Pseudo-Observations

Fumihiro Tajima, Keio University**; Kenichi Hayashi**, Keio University

Developing analyses of competing data is one of the central issues in biomedical statistics. Many approaches have been explored to overcome the essential incompleteness of the competing risk data. Andersen et al. (2003) proposed a method for such data based on pseudo-observations. The method allows the direct evaluation of the effect of covariates on the cause-specific cumulative incidence function (CIF) using an estimating equation with pseudo-observations. In the presence of multiple covariates such as patient's background, its selection is critical for effective estimation of the conditional CIF. This is an issue of variable selection and is essential to obtain an interpretable and predictable model.

In this study, we propose a variable selection criterion for competing risk data. This is derived by the asymptotically unbiased estimator of an expected quasi-likelihood based on the pseudo-observations. Numerical experiments show the performance of the proposed criterion outperforms the counterpart based on a naive extension of the AIC used for the ordinary context.

The Impact of Heterogeneity and Outliers on Flexible Shrinkage Estimators for Local Treatment Effects in Multi-Regional Clinical Trials

Naoki Isogawa, Pfizer R&D Japan; **Shintaro Hiro**, Pfizer R&D Japan; **Wang Wenjin**, Pfizer Inc.

In the development of new drugs, evaluation of similarity and difference in efficacy and safety among countries in multi-regional clinical trials (MRCT) has played an important role in regulatory decision making. A lot of evaluation strategies have been proposed and applied for the past two decades after ICH E5 became effective. Recently, ICH E17 was published and it is expected to have influence on the evaluation of treatment effects among countries in MRCT. A hurdle of the evaluation of local treatment effects in each country is the low precision in the estimation due to relative small sample sizes in some countries.

As one of the solutions to estimate local treatment effects, shrinkage estimators by borrowing overall results and similar countries results have been proposed by Chen et al. (2009) and Guo et al. (2015) respectively, but excessive shrinkage and borrowing may occur. This is due to a full exchangeability assumption of parameters. To overcome this, flexible shrinkage estimators (Neuenschwander et al., 2015; Gamalo-Siebers et al., 2016; Kaizer and Koopmeiners, 2018) have been introduced.

We evaluate the impact of the magnitude of the between-country heterogeneity and the outliers on the flexible shrinkage estimators through simulations under various scenarios and make some comments on the utilization of these shrinkage estimators.

Meta-Analysis and Matrix Decomposition for Pattern Extraction and Patient-Level Prediction of Adverse Events

Kentaro Matsuura, Johnson & Johnson; **Jun Tsuchida**, Tokyo University of Science; **Shuji Ando**, Tokyo University of Science; **Takashi Sozu**, Tokyo University of Science

In clinical studies, adverse events (AEs) occurring in each patient from the beginning to the end of the study are recorded. The primary purpose of assessing AEs is to determine which AE sets (patterns) are most likely to occur. A secondary purpose is to predict AEs at the patient level. The aim of this study was to develop a method which simultaneously serves both of these purposes. We constructed a statistical model including nonnegative matrix factorization by incorporating prior knowledge of AE-specific structures such as severity and drug mechanism of action. We also constructed a statistical model using a meta-analysis framework in order to be able to process data from multiple clinical studies because insufficient information is derived from a single trial. We used real data from three Phase III studies to evaluate the operating characteristics of the proposed method. The data included two mechanisms of action, five treatments, 3,317 patients, and 848 AE types. As a result, we extracted treatment-specific AE patterns such as “nausea and vomiting pattern”, “constipation pattern”, and “neuropathy pattern”, which coincided with background knowledge. Based on patient AE occurring within the first two treatment cycles, we quantitatively predicted the AEs most likely to occur in the latter two cycles.

Introducing the BGLIMM Procedure for Bayesian Generalized Linear Mixed Models

Amy Shi, SAS Institute; **Fang Chen**, SAS Institute

This talk introduces PROC BGLIMM, a newly released (SAS/STAT 15.1), high-performance, sampling-based procedure that provides full Bayesian inference for generalized linear mixed models. 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 or longitudinal data with repeated measurements. PROC BGLIMM provides convenient access, with improved performance, to Bayesian analysis of complex mixed models that you could previously perform with the MCMC procedure. This talk describes how to use the BGLIMM procedure for estimation, inference, and prediction.

Adaptive Power Prior for Sequential Clinical Trials - Application to Bringing Studies

Adrien Nigel Ollier, INSERM; **Satoshi Morita**, Kyoto University Graduate School of Medicine; **Moreno Ursino**, INSERM; **Sarah Zohar**, INSERM

This work is set in the context of bridging studies between two populations, for example Caucasians and Asians. These are early phase clinical trials designed to bridge and understand a possible “gap” between dose-response curves of different populations. In these cases, it might be desirable to share information using efficiently the available resources. We developed an adaptive power prior approach with a commensurate parameter for using historical or external information. It allows, at each stage, a full borrowing when data are not in conflict, no borrowing when data are in conflict or a ”tuned” borrowing when it is in between. We applied our method to bridging studies between Caucasians and Asians and we focus on sequential adaptive allocation design, although other settings can be used. We split the power prior parameter as a product of two terms computed adaptively. First, the effective sample size notion is used to set the maximum desirable amount of information to be shared from historical data at each stage of the trial; and then, in a sort of Empirical Bayes, a commensurability parameter is chosen using a measure of distribution distance. This approach avoids elicitation and computational issues regarding usual Empirical Bayes. We propose several versions of our method and we conducted an extensive simulation study evaluating the robustness and sensitivity to prior choices. Our simulations showed that our approach is

robust and can be easily adapted to each Bayesian analysis involving historical data, as well as in case of not sequential designs.

CS3	10:45AM - 12:00PM, August 30, 2019	Room 510, 5F

A Novel Bayesian Analysis of Dose-Response Relationship with Dynamic Generalized Linear Models in Oncology Phase I Study Using Power Priors to Incorporate Historical Data

Joji Mori, Eli Lilly Japan K.K.

Many study designs for dose escalation in oncology phase I studies assume toxicity probability monotonically increases with dose implicitly and explicitly. A logistic function used in CRM represents monotonicity well. However, recent molecules in immuno-oncology have not always shown monotonic dose-response. A new approach is necessary to understand dose-response of such molecules, and a dynamic generalized linear model (DGLM), originally developed in time series, can be used to estimate flexible dose-response, allowing for non-monotonicity. The precision of dose-response estimation is often problematic due to small sample size. It is effective to incorporate historical data and obtain precise estimates when an ongoing study mimics or is similar to past studies. Power priors have favorable features compared to other priors when incorporating historical data. To evaluate flexible dose-response precisely in small studies, we propose a novel Bayesian analysis to estimate dose-response with DGLM using power priors to incorporate historical data. We assumed to have data of both an ongoing study and a past study, and compared our analysis to frequentist inference based on the current study and a Bayesian analysis only using power priors to incorporate the past study. We carried out simulation studies to study how our analysis would improve estimation of toxicity probability at each dose compared to the others. Our simulation studies illustrated that both mean square errors and standard errors of the estimates for all doses were smallest in our analysis. We would say the more precise estimation of dose-response was possible in our analysis.

Adaptive Study Design using Model Based Dose Escalation with Two Pharmacodynamical Endpoints

Dion Chen, Janssen Research and Development

There are more and more new requests in drug development to make decision early. With small sample size and a lot of data usually in early phase studies, more statistical analysis needed for signal detection in safety and PK/PD. To identify the optimal dose regimen with two PD endpoints, a model-based dose escalation method is proposed for an early phase study.

Consider a new drug in early phase development, which targets CD38 which is highly expressed on PB-PC, that are considered as the primary source of auto-antibodies in systemic lupus erythematosus (SLE). The hypothesis is that depleting CD38 with the new drug in SLE will decrease auto-antibody levels and lead to clinical improvement. A phase 1b study is planned for SLE patients, the objectives is to identify dose and regimen key for efficacy, to understand the effects on autoantibodies to de-risk next step investment. Multiple doses are chosen in this study based on the preliminary results from the first-in-human (FIH) study in healthy subjects. The goals for dose escalation are aiming comparable and better effects than other B-cell targeting drugs and assessing safety and tolerability, maximizing dose range for immunological indications. An innovative adaptive study design is proposed in this study for does finding. The criteria for dose escalation are 70% of CD38 depletion at week 8 and 50% reduction of anti-dsDNA antibody at week 12. Using two PD endpoints, the dose escalation rules and sample size have been evaluated via simulation with model-based dose escalation method.

Bayesian Random-Effects Meta-Analysis of Phase I Dose-Finding Studies

Moreno Ursino, INSERM; **Christian Rver**, University Medical Center Gttingen; **Sarah Zohar**, INSERM; **Tim Friede**, University Medical Center Gttingen

Context: Phase I dose-finding studies aim at identifying the maximal tolerated dose (MTD). It is not uncommon that several dose-finding studies are conducted, although often with some variation in the administration mode or dose panel. For instance, Sorafenib (BAY 43-900) was used as monotherapy in at least 23 Phase I trials (clinicaltrials.gov). Since the toxicity may not be directly related to the specific indication, synthesizing the information from several studies might be worthwhile. However, this is rarely done in practice and only a fixed-effect meta-analysis framework was proposed to date.

Objective: To develop a random-effects meta-analysis methodology under Bayesian inference, without parametric dose-response model, to pool several Phase I trials and suggest the MTD.

Method: A curve free hierarchical model on logistic scale with random effects, accounting for heterogeneity between the trials, is used to model the probability of toxicity across the investigated doses. The Ornstein-Uhlenbeck Gaussian process is adopted for random effects structure. Prior distributions for the curve free model are based on a latent Gamma process, parametrized by “intercept”, “slope” and “variability”, and finally chosen using the effective sample size concept.

Results: The simulation study revealed that the percentage of correct MTD selection is 60-80%, when 10 trials are used for meta-analysis. The percentage decreases to 50-70% when using only 5 trials. The application to the Sorafenib example estimated the MTD at 600 mg.

Conclusions: Sharing information between phase I studies can improve the precision of MTD selection, at least when the number of trials is reasonably large."

Efficient, Doubly Robust Estimation of the Effect of Dose Switching for Switchers in a Randomised Clinical Trial

Kelly Van Lancker, Ghent University; **An Vandebosch**, Janssen R&D; **Stijn Vansteelandt**, Ghent University and London School of Hygiene and Tropical Medicine

The interpretation of intention-to-treat analyses of randomised clinical trials is often hindered as a result of noncompliance and treatment switching. This has recently given rise to a vigorous research activity on the identification and estimation of so-called estimands. Motivated by an ongoing clinical trial conducted by Janssen in which a flexible dosing regimen is compared to placebo, we evaluate how switchers in the treatment arm (i.e., patients who were switched to the higher dose) would have fared had they been kept on the low dose in order to understand whether flexible dosing is potentially beneficial for them. Comparing these patients’ responses with those of patients who stayed on the low dose does not likely entail a satisfactory evaluation because the latter patients are usually in a better health condition and the available information is too scarce to enable a reliable adjustment. In view of this, we will transport data from a fixed dosing trial that has been conducted concurrently on the same target, albeit not in an identical patient population. In particular, we will propose a doubly robust estimator, which relies on an outcome model and a propensity score model for the association between study and patient characteristics. The proposed estimator is easy to evaluate, asymptotically unbiased if either model is correctly specified and efficient (under the model defined by the restrictions on the propensity score) when both models are correctly specified. Monte Carlo simulations and application to a clinical trial conducted by Janssen demonstrated adequate performance.

Blinded Sample Size Re-estimation with Survival Data

Ryuji Uozumi, Kyoto University; **Shinjo Yada**, A2 Healthcare Corporation

Sample size estimation with survival data is generally conducted based on the expected number of events and the survival probabilities for each group in randomized double-blind controlled clinical trials. The number of events is mostly estimated under the proportional hazards assumption and each survival probability is frequently assumed to follow an exponential distribution for simplicity. We consider the setting in which misspecifying the survival distribution results in overestimation of the assumed event probability. In this talk, blinded sample size re-estimation for supplementing misspecification of the survival distribution will be discussed. The proposed procedure is performed using the blinded data pooled across groups by fitting parametric models. Results from simulation studies about the operational characteristics of each approach will be presented. Furthermore, real examples from clinical oncology are also illustrated.

Poster Session

PS	6:30 - 7:30PM, August 27, 2019	Swan, 1F

#50005: A Powerful Method to Meta-Analysis for Testing no Treatment Effects Cancelled

Kuang Fu Cheng, Asia University

In a meta-analysis of multiple trials, a fundamental problem is to test whether a new treatment of interest is better than the placebo or an active treatment. Because the data of a meta-analysis may be heterogeneous, fixed-effects approach and random-effects approach are often used in the literature. The fixed-effects approach assumes that the all trial effects are the same, while the random-effects approach usually assumes the effects are random and follow a normal distribution. Under the random-effects model, Han, B. & Eskin, E. (2011) suggested a test for testing the averaged effect and variance of the effects are both zero, that is, all treatment effects (log odds ratios) are zeros. In this paper, we propose a new approach for solving the same problem. The novel method has greater advantage in that no distributional assumption of normality is required. Furthermore, a simulation study indicated that it was more powerful than Han-Eskin test under wide range of simulation conditions.

#50029: Concentration-QTc Analysis for Phase 1 Studies without a Placebo Arm

Yasushi Orihashi, Tokai University School of Medicine; **Shoichi Ohwada**, Daiichi Sankyo Co., Ltd.; **Yuji Kumagai**, Kitasato University Hospital

BACKGROUND: Concentration-QTc (C-QTc) modeling has been increasingly used in phase 1 studies to evaluate the relationship between drug concentration and QTc interval. It is recommended that models include categorical time effects for placebo-controlled studies to account for circadian variations of the QTc interval. However, for studies without a placebo arm (e.g. oncology phase 1 studies), it is not obvious how circadian variations are accounted for.

METHODS: Simulations were conducted to investigate models with/without categorical time effects under various scenarios for typical phase 1 studies without a placebo arm. We evaluated the bias, coverage probability, standard error of estimates, and probability concluding no QTc prolongation, i.e. false negative rate for drugs that prolong the QTc interval and true negative rate for drugs that don't.

RESULTS: In most of the scenarios, the models without categorical time effects caused bias in estimating the C-QTc relationship and could not keep the nominal false negative rate, whereas the models including categorical time effects provided unbiased estimation. In addition, the QTc interval data on a time-matched baseline day or a day after multiple dosing improved the precision in estimating the C-QTc relationship. The QTc interval data at lower doses than therapeutic doses improved the precision as well, even if the sample size in the lower doses is small. Further, increasing the number of subjects in the therapeutic doses contributed to increasing the precision and the true negative rate.

CONCLUSIONS: We recommend that C-QTc models include categorical time effects for studies without a placebo arm, too.

#50030: Model Selection for Semiparametric Marginal Mean Regression Accounting for Within-Cluster Subsampling Variability and Informative Cluster Size

Chung Wei Shen, National Chung Cheng University; **Yi Hau Chen**, Institute of Statistical Science, Academia Sinica

We propose a model selection criterion for semiparametric marginal mean regression based on generalized estimating equations. The work is motivated by a longitudinal study on the physical frailty outcome in the elderly, where the cluster size, that is, the number of the observed outcomes in each subject, is "informative" in the sense that it is related to the frailty outcome itself. The new proposal, called Resampling Cluster Information Criterion (RCIC), is based on the resampling idea utilized in the within-cluster resampling method (Hoffman, Sen, and Weinberg, 2001, Biometrika 88, 1121-1134) and accommodates informative cluster size. The implementation of RCIC, however, is free of performing actual resampling of the data and hence is computationally convenient. Compared with the existing model selection methods for marginal mean regression, the RCIC method incorporates an additional component accounting for variability of the

model over within-cluster subsampling, and leads to remarkable improvements in selecting the correct model, regardless of whether the cluster size is informative or not. Applying the RCIC method to the longitudinal frailty study, we identify being female, old age, low income and life satisfaction, and chronic health conditions as significant risk factors for physical frailty in the elderly.

#50031: Bayesian Flexible Modeling the Odds under Case II Interval-Censored Data

Li-Chu Chien, Kaohsiung Medical University; **Yuh-Jenn Wu**, Chung Yuan Christian University; **Wei-Quan Fang**, Academia Sinica; **Li-Hsueh Cheng**, Chung Yuan Christian University

Interval-censored time-to-event data often arise in medical, biological, epidemiological or health studies. However, most of existing interval-censored survival analysis approaches often encounter difficulties such as heavy computational complexity or non-proportionality of hazard rates due to the complex time-to-event data structures. To address these difficulties, in this study, we use the Bayesian nonparametric algorithm to estimate the odds based on case II interval-censored data. We introduce Bernstein priors for modeling the odds and propose an easy sampling way to study the posterior distributions through the Markov chain Monte Carlo algorithms. We also investigate asymptotic properties of the posterior distributions. The simulated examples show that the proposed approach works well in the cases considered. The practicality of the proposed method is illustrated with the hemophilia study data analyzed.

#50032: Role of Baseline Covariates in ex-Vivo Bioassay for the Assessment of Intrasubject Parallelism

Hideaki Uehara, Tsumura & Co.; **Kazuko Satoh**, Tsumura & Co.; **Nagisa Komokata**, Tsumura & Co.; **Kazuo Ogawa**, Tsumura & Co.

In the ex vivo dose-response (D-R) experiment which uses the animal as the experimental Block, we often observe a variety of D-R curves which we deem as the intersubject heterogeneities. To some extent, we may explain this phenomenon by the varied responsiveness of specimens, which we may observe at baseline in two different manners: the spontaneous reactions with and without the administration of standard stimulant. We can deem the former as the predictor or the surrogate of the maximal response (plateau), and the latter of the minimal response (nadir) in that specimen.

Meanwhile, the between-substance similarities of D-R curves are crucial for the relative potency estimation. In the case of the parallel line bioassay, for example, we need to demonstrate the parallelism between D-R slopes. When using multiple animals, it is also crucial to confirm the parallelism within each animal, as well as per population average (Uehara et al. 2016a). In this regards the two variables above should be useful.

In this paper, we discuss the appropriate use of these covariates in the assessment of intrasubject parallelism (Uehara 2018). To improve its efficiency, we modeled the covariates' influence over the D-R curves to extract the heterogeneity which the covariates cannot explain. We present an example which uses the proposed method along with some simulation study results. We also discuss the issues related to the similarities in asymptotes, which we implicitly assume but are out of scope in the parallel line assay.

#50042: Log-rank Test and Its Handicap Procedure Using Computational Algebraic Statistics

Kotaro Mizuma, Osaka University; **Tomoyuki Sugimoto**, Shiga University

In survival analysis, log-rank test is commonly used to test whether two or more survival curves are different. The test procedure based on asymptotic theory often becomes inappropriate in imbalanced datasets. Alternatively, an exact inference can be used in this setting. This approach doesn't utilize asymptotic theory and need a large sample assumption. However, this method also has a flaw; the exact method must calculate all the candidates that can be obtained, so it is problematic in computational complexity. We propose a new test based on exact inference that uses computational algebraic statistics. It is infeasible to calculate a p-value using exact inference since sample spaces tend to be complicated and large. On the other hand, our method calculates using a Markov Chain Monte Carlo (MCMC) approach. In this approach, we construct a Markov chain that converges to the exact probability and then estimate the p-value from the Markov chain obtained. Our method can be applied even in imbalanced datasets and complicated datasets. Also, our method can be applied to a handicap procedure for non-inferiority testing based on log-rank test,

since an MCMC method doesn't need all the information; i.e., an MCMC can be constructed by only a proportional term of the exact probability. We'll also show some simulation results and its performance.

#50047: Adaptive Randomization for Multiarm Survival Clinical Trials Using Short-Term Response Information

Yu Mei Chang, Tunghai University

The main goal of cancer clinical trials is to prolong the survival time for cancer patients. However, it takes a long time to observe the survival endpoint. The information about a short-term response is often quickly available during or shortly after treatment, and this short-term response is a good predicator for long-term survival. In this paper, we proposed a randomization design for multiarm clinical trials, in which the primary endpoint is survival but incorporate the information about the short-term patient response in order to implement a more effective adaptive randomization. The short-term response and the long-term survival are connected through a Bayesian mixture model. The adaptive deign first uses prior clinical information of the experimental treatments and then updates the model dynamically according to information accumulated to guide and modify the ongoing trial such that patients have a greater chance of being allocated to more superior treatments. We also use the posterior distributions to set up early stopping criteria and implement an outcome-adaptive patient allocation algorithm. We conduct simulation studies under different scenarios to examine the finite sample performance of our proposed method.

#50049: A Robust Association Test with Multiple Genetic Variants and Covariates

Jen-Yu Lee, Feng Chia University; **Kuang-Fu Cheng**, Asia University and Taipei Medical University

Since the sequencing technology of discoveries and analyses of genetic variants at a gene or exome have a great stride. To study the association between disease and genetic variants become a feasible issue for now. Some powerful and well-known association tests have been proposed to test the exist of significance relationship with a group of genes which may associate with a disease of interest. However, some challenges still remain, for examples, the sample size, number of causal and non-causal variants, and size and direction of effect, all affect the performance of testing power. Recently, a powerful method named TREM which derived from a random-effects model and allows for missing genotypes was proposed. This method is less sensitive to the inclusion of non-causal rare variants and/or low effect common variants or to the presence of missing genotypes, and also when the effects were more consistent in the same direction. The drawback of TREM is also obviously when effects are not in the same direction. Here, we proposed a novel testing method which is robustness when these situations met. Our simulation results show that type I error rates and power of several competing tests under various conditions. The proposed method has stable type I error rate and better power performance in most scenarios.

#50055: Bayesian Model Selection on the Structural Equation Model: An Application to a Longitudinal Myopia Trial

Yi-Fu Wang, National Chung Cheng University; **Tsai-Hung Fan**, National Central University

Structural equation models (SEMs) have been extensively used in behavioral, social, and psychological research to model relations between latent variables and observations. This article introduces a Bayesian model selection problem to the SEMs. To put the reasonable mixture prior on the specific parameter which describes the doubting relationship in the SEMs, the model posterior probability can be computed via the MCMC iterations and viewed as a Bayesian model selection criterion. An advantage of the method using mixture priors is that it can automatically identify the predictors having non-zero fixed effect coefficients or non-zero random effects variance in the MCMC procedure. The proposed methodologies are illustrated through a simulation study. Specifically, we will focus on the multidimensional longitudinal myopia data to reduce the dimensionality of the parameter space and to select the simpler model.

#50056: Statistical Approach with Right-Censored Survival Data for Design and Evaluation in the Multiregional Clinical Trial

Yu-Chieh Zheng, National Health Research Institutes; **Yuh-Jenn Wu**, Chung Yuan Christian University,

Hsiao-Hui Tsou, National Health Research Institutes and China Medical University

Multiregional clinical trials are conducted increasingly for convenience and safety of patient under the same protocol in many different areas at the same time. Overall survival is very important endpoint, so many multiregional clinical trials are conducted with the survival endpoint. The right censored survival situation is usually happened. After showing the overall efficacy in all participated regions, the local regulatory may require the sponsors to provide evidence of consistency in the treatment between the overall patient and local region. It is done for claiming the drug' efficacy and getting drug approval in the local region. In this paper, we established a random effect model of hazard rate with right censored case to design and evaluate treatment effect in the multiregional clinical trial. For simplify, we consider constant hazard rates for test and control groups. The difference of hazard rates for control and test groups based our approach are used to test hypothesis in the design stage. We provide the method of sample size determination. Assurance probability in our approach is to evaluate positive regional difference of hazard rates and positive overall treatment effect simultaneously. Furthermore, we discuss the impact of design parameters, such as follow up time, regional hazard rates for a test and a control groups, regional proportion on total sample size, by simulation. An estimated hazard rate is calculated by the maximum likelihood estimation.

#50065: Using PMDA Drug Adverse Event Report Database, Study on Collective Background of Adverse Events

Shoko Kamiya, Keio Research Institute at SFC; **Michiko Watanabe**, Keio University; **Keita Yamaguchi**, Keio University

Objective: This study examines the collective background and characteristics of patients with adverse events caused by total cold medicines.

Methods: Latent class analysis is performed on 990 cases reported between April 2004 and June 2015 using PMDA's database (JADER). In addition, the signal detection is implemented using the same DB, and it evaluated about the relationship with the result of latent class analysis.

Results: As a result of Latent class analysis, the target population was divided into three classes. Class 1 was a group that has no primary disease or medication. Class 1 was named "health group", accounting for 53.7% of the whole. The special adverse event was an immune system disease. Class 2 was interested in self-treatment, accounting for 33.2% of the whole, and was named "self-medication group". The special adverse event was a serious skin disease. Class 3 was 13.1% of the whole, 90% of this population is over 60 years old, and most have primary disease and medication, so it was designated as "upper-aged outpatient group". The special adverse events were lung disease and nervous system disorders. Moreover, it became clear that the signal by signal detection was detected in the high occurrence group (class) in all the adverse events of medicines, and the collective background could be related.

Conclusions: Latent class analysis could reveal the collective background and characteristics of adverse event occurrence. This research is applicable to other medicines and is expected to contribute as a new application of JADER.

#50071: The Use of Maximum a Posteriori Estimation for Selecting Dose in Phase I Clinical Trials

Wen-Jin Guo, National Health Research Institutes; **Chieh Chiang**, National Health Research Institutes; **Chin-Fu Hsiao**, National Health Research Institutes

The purpose of phase I clinical trials is to find the maximum tolerated dose of an experimental drug. Traditionally, the 3+3 design and the continual reassessment method are the most common assessment approaches for the trials in cancer. In 2017, Guo et al. proposed a Bayesian design, called modified toxicity probability interval-2 (mTPI-2) method, to improve the procedure of dose selection. We extended the concept from mTPI-2 to the use of maximum a posteriori estimation. The new method is simpler than mTPI-2. A dose finding table is provided in our study despite the fact that users can easily select the experimental dose level through a calculator. A numerical study is provided for comparing our proposed method with mTPI-2.

#50075: Bioequivalence Assessment between Sugar-coated and Film-coated Eperisone Tablets using Reference Replicated Crossover Study for Highly Variable Drug

In-Hwan Baek, Kyungsung University

Bioequivalence assessment of highly variable drugs is complicated because the large number of volunteers required and the risk of erroneously rejecting BE is remained. In this study, we analyzed the within-subject variability (CVwR) of sugar-coated eperisone tablets (reference formulation) for highly variable drugs and to conduct bioequivalence study of sugar-coated and film-coated tablets (test formulation) of eperisone hydrochloride 50 mg in healthy Korean volunteers by reference-replicated crossover study. Thirty-six healthy Korean male volunteers were recruited, and 33 subjects completed the study. A randomized, single-dose, open-label, three-way, three-sequence, reference formulation-replicated, crossover bioequivalence study was conducted to determine the bioequivalence of eperisone. Noncompartmental pharmacokinetic analyses were conducted using Phoenix WinNonlin. CVwR was calculated by using analysis of variance (ANOVA) on the reference data only, with sequence, subject-within sequence, and study period as fixed effects. The point estimates and 90% confidence intervals (CIs) for the test/reference geometric mean ratio (GMR) were calculated for AUCt and Cmax. and presented as least-squares means. The CVwR of eperisone reference product was 33.17 % for AUC12 h and 50.21 % for Cmax. The acceptance limit for Cmax was scaled to 0.69841.4319 according to CVwR. The 90% CIs for the test/reference geometric mean ratio were 0.82751.1692 for AUC12 h and 0.75871.1652 for Cmax, which were within the accepted bioequivalence limits. Therefore, the newly developed film-coated tablet is interchangeable with the original sugar-coated tablet of eperisone.

#50081: Clustering-based Basket Trial Design for Assessing Heterogeneity of Treatment Effect among Strata

Ryo Sadachi, The University of Tokyo; **Akihiro Hirakawa**, The University of Tokyo

A basket trial in oncology includes in patients bearing different cancer types with a common biomarker and evaluate treatment effect in each stratum. In the basket trial, assessing heterogeneity of treatment effect among multiple strata is challenging. A Bayesian hierarchical modeling (BHM) has an attractive feature of stabilizing the estimate of treatment effect in each stratum, but bears the risk of arbitrariness to determine the degree of shrinkage and the similarity of treatment effect between strata.

In this study, we develop a new method for assessing the heterogeneity of treatment effect and estimating treatment effect in each stratum. The proposed method quantifies the similarity of treatment effect between two strata based on the Jensen-Shannon divergences and groups the strata by applying the aggregative hierarchical clustering for Jensen-Shannon divergence. Subsequently, we evaluate treatment effect in each cluster using the binomial exact test. Via simulations, the proposed method determines the required sample size in each stratum and the threshold for determining the optimal number of clusters while maintaining the family wise error rate and marginal power at a certain level. To compare to independent method that evaluates treatment effect in each stratum independently, we examine the utility of the proposed method through the simulation studies.

#50087: A Joint Modeling Approach for Predictions of Survival Based on Tumor Dynamics and New Lesions in EGFR Mutation-Positive Non-Small Cell Lung Cancer Patients Treated with Gefitinib or Carboplatin and Paclitaxel

Mario Nagase, AstraZeneca; **Katsuomi Ichikawa**, AstraZeneca K.K.; **Jacqueline Buros-Novik**, Icahn School of Medicine at Mount Sinai; **Diansong Zhou**, AstraZeneca; **James Dunyak**, AstraZeneca; **Nidal Al-Huniti**, AstraZeneca

BACKGROUND: Tumor burden and its growth rate have a great impact on predicting progression-free survival (PFS) and overall survival (OS) in oncology clinical trials. A statistical framework to predict both survival probabilities at progression in Non-Small Cell Lung Cancer (NSCLC) patients is needed.

METHODS: We applied a joint model which incorporates longitudinal tumor burden, appearance of new lesions, and changes in therapy following progression in NSCLC patients, to interrogate the components of RECIST and to predict PFS. A nonlinear exponential growth model for time-dependent tumor growth is assumed and the patient-treatment-course-specific parameters are informed by a subset of the patient-treatment-course-specific covariates. The baseline hazard was estimated using a cubic spline with 6 knots, and the hazard to estimate patient-specific survival was numerically integrated using

Gauss-Kronrad quadrature with 13 quadrature points. Parameters were estimated from the posterior distribution and realized with R and STAN.

RESULTS: The model was fitted to data from 434 NSCLC patients treated with gefitinib or carboplatin+paclitaxel ('IPASS', NCT00322452), and further evaluated on 102 EGFR+ NSCLC patients treated with gefitinib ('IFUM', NCT01203917). The prediction accuracy for observed data and recapitulates RECIST outcomes was nearly 90%. The sum-of-longest-diameters (SLD) and new lesions were associated with OS (hazard ratios: 1.96 and 6.40), while the rate of change in SLD showed little association with OS (hazard ratio 1.00).

CONCLUSION: The Bayesian joint model accurately recapitulates the RECIST-based outcomes and generates well calibrated predictions of survival. It offers insights regarding the relative predictive value of the components of RESICT in NSCLC.

#50091: Non-Asymptotic Properties and Behaviors for Random-Effects Meta-Analyses When the Number of Studies Is Small

Keisuke Hanada Kagoshima University; **Tomoyuki Sugimoto**, Shiga University

Meta-analysis, which consolidates multiple studies of the same treatment or problem, is frequently used in medicine and other science fields. Random effects model is often used in meta-analysis because they can perform analysis taking into account heterogeneity between studies. However, in random effects meta-analysis, asymptotic normality is used to infer the processing effects, so that it is difficult to perform more accurate estimation when the number of studies is smaller. In this study, we investigate the properties and behaviors of the estimator for the random effects meta-analysis using an exact distribution of the heterogeneity between studies and discuss how accurately the method based on the asymptotic approximation can be analyzed. We discuss the non-asymptotic method for the random effects model, and show that the Pvalue from the conventional method (e.g., DerSimonian-Laird) tends to be smaller relative to an exact method, especially when the number of studies is small, using the numerical simulation. We also present an exact construction for the confidence interval of the estimator, because the conventional method makes the interval narrowed. Therefore, when the number of studies is smaller, it is useful to construct the exact distribution and confidence interval for the random effects meta-analysis.

#50098: Comparison of Hazards in Two-Arm Trials with Exponential Distributed Outcomes from the Bayesian Viewpoint

Masaaki Doi, Kyoto University; **Yohei Kawasaki**, Chiba University Hospital

We compare the hazards of two exponential distributions based on the Bayesian posterior probability of some hypotheses being true, incorporating historical information by using the conditional power prior. We investigate the Bayesian evidences of superiority, non-inferiority, and equivalence of hazards. In each case, we derive simple and easy-to-calculate expressions for Bayesian posterior probabilities. Furthermore, for the case without censoring, we show clear relationships between the Bayesian posterior probabilities with some non-informative priors and p-values of the "exact" F-test. These relationships can be seen as the Bayesian- frequentist reconciliations. We use Monte Carlo simulations to evaluate the operating characteristics and analyze data on two melanoma clinical trials.

#50110: Comparison of Bayesian Equivalency Methods for Two Binomial Outcomes Using Bayesian Index

Yohei Kawasaki, Chiba University; **Yosuke Inaba**, Chiba University; **Hideki Hanaoka**, Chiba University; **Etsuo Miyaoka**, Tokyo University of Science

In clinical trials, it is often necessary to perform an equivalence study. The equivalence study requires actively denoting equivalence between two different drugs or treatments. Since it is not possible to assert equivalence that is not rejected by a superiority test, statistical methods known as equivalency tests have been suggested. These methods for equivalency tests are based on the frequency framework; however, there are few such methods in the Bayesian framework. Hence, this study proposes a new index that shows the equivalency of binomial proportions, which is constructed based on the Bayesian framework. In this study, we provide some methods for calculating the index and compare the probabilities that have been

calculated by these two calculation methods. Moreover, we apply this index to the results of actual clinical trials to demonstrate the utility of the index.

#50137: G-estimation of Structural Nested Mean Models for Interval-Censored Data Using Pseudo-Observations

Shiro Tanaka, Kyoto University

Standard survival analysis assumes that time-to-event is observed exactly or right-censored, meaning that it is only known that the event occurred after the last observation. In some situations, however, time-to-event may only be known to exist within an interval of time. These data are interval-censored. A typical example is incidence of diabetes retinopathy between fundus examinations at pre-scheduled visits. In this presentation, I propose structural nested mean models and g-estimation for causal inference with interval-censored data. Inference is implemented by g-estimation using pseudo-observations, a technique to handle censoring. The finite-sample performance of the proposed estimators in simulated datasets will be presented.

#50138: Bayesian Evidence Synthesis and Assessment Techniques across Longitudinal Time Points

Airi Takagi, Tohoku University; **Takuhiro Yamaguchi**, Tohoku University Graduate School of Medicine

Mixed treatment comparisons are increasingly being used to evaluate clinical effectiveness. Different studies often have differing durations and various time points; therefore, several methods for considering multiple time points have been proposed.

These methods focus on the summary of outcome measure values for comparison between treatment groups or studies. Change in the disease progression and healing rates in each study affect the outcome measure values and lead to between-study variability, therefore kinetic analysis methods are considered useful; however, this has seldom been addressed.

In this work, we extend the current methods and propose a Bayesian estimation using the kinetic analysis methods of reversible reaction to synthesize longitudinal summary level data of continuous outcomes and discuss whether the indicator using the kinetic constant is a useful comparison of effectiveness. The rate law equation assumes that the time rate of reduction in the observed mean y changes linearly with respect to quantity y, and we obtain the exponential function by integration. Our approach includes terms to consider differences in both the disease progression and healing rates caused by differences in study characteristics or design. Correlation across multiple time points can be considered while applying the model to longitudinal summary level data.

We conducted simulation studies to assess the numerical performance of our proposed method. Methodological developments were illustrated using summary level data of HAMD score measurements in the treatment of depression, HbA1c in the treatment of type 2 diabetes and some other outcome measure.

#50146: Patient Subtypes Associated with Medication Persistence Using Latent Class Analysis

Shiori Nishimura, Keio University; **Michiko Watanabe**, Keio University; **Keita Yamauchi**, Keio University

Background: Non-persistence to medication is one of the primary reasons for suboptimal outcome in mental illness. Few studies evaluated medication adherence and persistence among patients with bipolar disorder in real-life setting. Our study aims to assess the association between patient subtypes and medication persistence using administrative claims database.

Methods: Using employee's health insurance claims database, all patients aged 20 years or older from January 2005 to October 2016 who had at least 1 record for bipolar disorder (ICD10 code: F31) were included. All patients were enrolled in a health insurance for over 3 years following the index date. We conducted two latent class analysis (LCA) to classify patients characteristics (i.e. medications, comorbidities). Bayesian information criterion(BIC) was used for model selection.

We also assessed the association between persistence and patient subtypes. Persistence was defined as

the duration prescribed medications during follow-up.

Results: A total of 3,741 patients met the inclusion criteria. Six classes were identified with different medication patterns(i.e. All low (28%), Lower Anxiolytics(21%), Lower Antipsychotics(20%), Multiple Medication and Higher Antidepressants(17%), Lower Antidepressants and Higher Antipsychotics(8%), Multiple Medication and Higher Hypnotics sedatives) . Three distinct classes with different comorbidity patterns were also identified(i.e. All low(71%), Higher Lifestyle-related Disease(21%), Lower Lifestyle-related Disease(8%)). Subtypes including multiple medication were more likely to be persistent.

Conclusion: Experience of multiple medication associates with higher persistence in real-life setting. Finding also suggests that interventions focused on patients who treated with simple medication may help improvement of persistence.

#50150: A Robust Covariate Selection Method for the Limited Sampling Design in Population Pharmacokinetic Analysis

Asuka Nemoto, Teikyo University Graduate School of Public Health

The selection of covariates explaining between subject variability in the pharmacokinetic (PK) parameters is routinely conducted in population pharmacokinetic (PPK) model building. To determine whether a candidate covariate is included into or excluded from the model, the likelihood ratio test is typically conducted. There is a time region that is 'information-rich' with respect to each PK model parameters, as can be shown by the plot of the partial derivative of the predicted concentrations to each of PK parameters vs time. We call 'limited sampling design' to the study without enough observations in an 'information-rich' time region for some reasons, e.g., studies targeted the pediatric patients, or a PPK study planed as an add-on study. The aim of this study was to propose a new method for selecting a covariate robust for a limited sampling design.

Wang (2007) provided the derivation of approximated two times of the negative logarithm of the likelihood (-2LL) using the first-order conditional method. With our proposed equation, a component of -2LL for arbitrary times other than actual sampling times for a subject i, can be calculated.

We hypothesized that this approach provides a more robust method for selecting a covariate in a limited sampling design than the conventional method. Monte Carlo simulations and estimations were used to evaluate this hypothesis in some limited sampling designs. Preliminary results suggest that the proposed method can substantially improve the power for selecting a covariate of the clearance when there is no model miss-specification and little measurement error.

#50163: Mediation and interaction of age, follicle stimulating hormone (FSH) and anti-müllerian hormone (AMH) on in vitro fertilization pregnancy

Han-Chih Hsieh, Institute of Statistical Science, Academia Sinica; **Jia-Ying Su,** Institute of Statistical Science, Academia Sinica; **Yen-Tsung Huang**, Institute of Statistical Science, Academia Sinica; **Shunping Wang**, Women and Infants Hospital in Rhode Island

Both follicle stimulating hormone (FSH) and anti-m¨ullerian hormone (AMH) are widely used to assess the ovarian reserve in women for in vitro fertilization (IVF). However, studies also showed that both AMH and FSH are significantly associated with age: as age increases, AMH decreases and FSH increases. This study aims to investigate the mechanism of age effect on IVF live birth rate, particularly through mediation and interaction by AMH and FSH. We conducted a retrospective cohort study of 13970 IVF cycles collected by eIVF from 2010 to 2016. A series of logistic mixed models were used to estimate the association of live birth and AMH (or FSH). The mediation effects and proportion mediated, were quantified by our previously proposed mediation analyses. We further investigated the FSH-modified mediation effects on live birth rate through AMH, accounting for the nonlinear age effect. Our analyses showed that age effect on live birth was mediated more by AMH than by FSH (18 vs. 6 %). The mediation effect through AMH can be further modified by FSH level regardless of women's age. In summary, mediation model provides a new perspective elucidating the mechanism of IVF successful rate by age.

#50164: Semiparametric Causal Mediation Modeling of Semi-Competing Risks

Ju-Sheng Hong, Institute of Statistical Science, Academia Sinica**; Shu-Hsien Cho**, Institute of Statistical Science, Academia Sinica**; Yen-Tsung Huang** Institute of Statistical Science, Academia Sinica

Semi-competing risks are frequently encountered in biomedical research such as clinical trials in which a primary outcome (e.g. overall survival) may censor an intermediate event (e.g. cancer recurrence) but not vice versa. We propose a semiparametric approach formulating the semi-competing risks as a causal mediation problem. We develop a mediation model with the intermediate and primary events, respectively as the mediator and the outcome. Counting process-based indirect and direct effects, respectively, are defined as an effect of an intervention on the primary outcome mediated through the intermediate event, and that not mediated through the intermediate event. We construct Breslow-type hazard estimators for direct and indirect effects, with time-varying weights. Asymptotic properties are established for the proposed estimators. Using simulations, we evaluate the finite-sample performance of the proposed estimators. The utility of our proposed methods is illustrated in a hepatitis study of liver cancer survival.

#50165: A Novel Extension of Keyboard Design: MT-Keyboard with Multiple Toxicity Constraints

Fangrong Yan, China Pharmaceutical University; **Ying Yuan**, University of Texas MD Anderson Cancer Center; **Liyun Jiang**, China Pharmaceutical University

In oncology, the toxicity of immunotherapy is often mild to moderate (grade 1-2), which brought up a challenge to the traditional designs with DLT, typically defined as grade 3 or higher toxicity. Under this definition, the toxicity profile observed in immunotherapy trial may be sparse, which may cause risky decisions during the dose decision-making period. In that case, we propose a novel extension of Keyboard design that accounts for multiple toxicity constraints, referred as MT-keyboard. Similar to the 3+3 design, MT-keyboard is easy to implement because its dose escalation/de-escalation rule can be pre-tabulated before the onset of trial, but has higher accuracy in identifying the MTD. Compared to the existing model-based method MC-CRM that deals with low grade toxicity, MT-keyboard has comparable accuracy to identify the MTD, but is easier to implement and has better performance in terms of safety, e.g., lower percentage of overdosing selection and overdosing patients. Simulation results show that MT-keyboard has high accuracy to identify the MTD without losing safety and thus has good potential in immunotherapy drug trials with low grade toxicity.

#50166: Causal Mediation of Chronic Hepatitis B or C on Mortality through Liver Cancer Incidence

Yi-Ting Huang, Institute of Statistical Science, Academia Sinica**; Yen-Tsung Huang**, Institute of Statistical Science, Academia Sinica

Hepatitis B and C viruses globally cause about 887,000 and 399,000 deaths per year, respectively. Although liver cancer is a common cause of death in carriers of hepatitis B virus (HBV) and hepatitis C virus (HCV), HCV infection has been construed as a systematic disease with severe extrahepatic manifestations. In this study, we aim to investigate the natural course of HBV or HCV infection using causal mediation modeling. A population-based cohort study (Risk Evaluation of Viral Load Elevation and Associated Liver, REVEAL) was conducted in Taiwan, in which 23,820 individuals recruited during 1991 and 1992 were followed prospectively until 2008. Chronic hepatitis B and C infections respectively were characterized by hepatitis B surface antigen and antibodies against hepatitis C virus in serum samples. The dataset was linked to the National Cancer Registry and the Death Certification System in Taiwan to determine liver cancer incidence and mortality. A semiparametric causal mediation model was implemented to estimate the causal mediation effect of hepatitis on mortality in relation to liver cancer incidence. Overall, we found that the effect of HBV infection on mortality was mostly through liver cancer incidence while the mortality of HCV carriers can be mediated through liver cancer or other diseases. The pattern of causal mediation also varied by age and gender. In conclusion, our mediation analyses demonstrate different natural histories of HBV and HCV infections, which may guide either therapeutic or policy interventions to reduce preventable deaths in this high-risk population.

#50167: Enterprise Investment Selection for the Kickstarter Projects

Yu-Jie Huang, National Sun Yat-sen University

Kickstarter is a platform that innovative projects can use to acquire funds through crowdfunding. In recent years, many novel products were successfully developed after obtaining funding and promotion on this platform. This study aims to provide an innovative Kickstarter technology project recommendation model to guide investment strategies. We use word2vec and Bag of Words to extract the text information from the website, then combine features to establish inter-project similarity scores in a content-based approach.

Some companies aim to discover business concepts with commercial value or find innovative opportunities for investment through this platform. However, there are hundreds of thousands of fundraising projects on the Kickstarter platform, which are usually represented by a large amount of text. It is a challenge to investigate the projects manually one by one. Our model result is 18.4% higher than the random recommendation.

Regulatory Symposium

Regulatory Symposium	3:30 - 5:00PM, August 29, 2019	Sakura, 1F

RS03: ICH-E17: How to Implement Multi-Regional Clinical Trials based on the Guidance?

TBD

Osamu Komiyama, Pfizer Japan Inc.

TBD

William W. Wang, Merck Sharp & Dohme Corp.

TBD

Hiroyuki Sato, Pharmaceutical and Medical Devices Agency

Regulatory Symposium	3:30 - 5:00PM, August 29, 2019	Room 510, 5F

RS01: Harmonization of Model-Informed Drug Development Approaches in Regulatory Review and Decisions

Model Informed Drug Development Good Practices: An Industry Perspective

Scott F. Marshall, Pfizer R&D LTD

In recent years, there has been increased interest in the development of good practices with respect to model-informed drug (discovery) and development (MIDD /MID3) with the aim of improving its implementation, standardization, and acceptance within drug development and regulatory review (1-4). Specifically, an EFPIA consortium has proposed good practices to help to reduce the heterogeneity in both the general quality and content of MID3 documentation in regulatory submissions (3). In parallel good practice proposals for specific applications e.g. population pharmacokinetics (5,6) exposure response (7) physiologically based PK (8) and QSP (9) and model evaluation (10) have evolved from individual company practices and professional body working parties.

While joint clinical pharmacology and statistical community interactions has, amongst other things, led to the sharing of good practices (11), the continued development of a common understanding of terminology and assumptions will be essential to the successful interplay between of MIDD and innovative statistical study designs.

Motivational publications, guideline development and policy evolution such as those proposed in this session by the FDA (12,13) and other global regulatory bodies (14,15,16) ; along with the future prospect of global harmonisation to facilitate the consistency with respect to both the expectation and acceptance of these approaches are welcomed by the industry.

This presentation will overview these developments from an pharmaceutical industry perspective, explore with reference to the results of a recent pan industry-regulatory survey the importance of the education and orientation of industry and regulatory decision makers to these approaches (17).

Regulatory Symposium	8:45 - 10:15AM, August 30, 2019	Room 510, 5F

RS04: ICH-E9(R1)- Guidance Concept and Implementation

TBD
Robert J. Hemmings, Consilium Salmonson and Hemmings

TBD
from Industry

TBD
from PMDA

Pre-Conference Short Courses

Short Couse	9:30AM - 12:45PM, August 26, 2019	Room #663, 5F

A. Missing Data Analysis in Clinical Trials Using SAS®

Instructor(s): **G. Frank Liu**, Merck, Sharp & Dohme, Corp; **Fang Chen**, SAS Institute Inc.

Missing data are inevitable and post many issues and challenge in analysis for clinical trials. Despite a great amount of research has been devoted to this topic, properly handling missing data in clinical trials remains complex. Conventionally, under the missing at random (MAR) assumption, we often use maximum likelihood or multiple imputation based methods for inferences. However, the MAR assumption is unverifiable. More critically, the estimand under MAR is hypothetical as indicated in the recent ICH E9 (R1) addendum and has been considered as overly-simplistic and unrealistic. Both regulatory agencies and industry sponsors have been seeking alternative approaches to handle missing data in clinical trials under missing not at random (MNAR) assumption.

This half-day tutorial is intended to cover issues of missing data in clinical trials including various methods and how to carry out the analyses using SAS software. The tutorial begins with an overview of missing data issues, and concepts and strategies as proposed by ICH E9 (R1) addendum. Then we will review traditional missing data handling methods such as maximum likelihood methods, multiple imputation, generalized estimation equation approaches, and Bayesian methods. The rest of the course is devoted to more recently-developed methods, such as sensitivity analysis to assess robustness, control-based imputation, control-based mean imputation, trimmed mean and tipping point analysis. Real clinical trial examples will be presented for illustration with implementation of the analysis using SAS/STAT software, including PROC MIXED, PROC MI, PROC MIANALYZE, PROC GEE, and PROC MCMC.

Outline

1. Review missing data issues and ICH E-9 (R1)
   Missing Data Issues in Clinical Trials
   Review of ICH E9 (R1): Estimand and 5 Strategies

2. Traditional methods for missing data
   Mixed-effects Model Repeated Measure (MMRM)
   Constrained Longitudinal Data Analysis (cLDA)
   GEE and wGEE
   Multiple Imputation (MI)

3. Recently-developed methods under MNAR
   Bayesian Methods
   Control-based Imputation, Control-based Mean Imputation
   Tipping Point Methods
   Trimmed Mean Analysis

Short Couse	9:30AM - 12:45PM, August 26, 2019	Room #664, 5F

B. Confirmatory Adaptive Designs with Multiple Objectives

Instructor(s): **Frank Bretz**, Novartis AG; **Franz König**, Medical University of Vienna; **H.M. James Hung**, Food and Drug Administration; **Sue-Jane Wang**, Food and Drug Administration; **Martin Posch**, Medical University of Vienna

Adaptive (flexible) designs allow for mid-course design adaptations based on interim data without compromising the overall type I error rate. Examples of design adaptations are the adjustment of sample sizes or the number and timing of interim analyses. These design parameters may be adapted depending on interim estimates of the variance, the treatment effect and safety parameters. An important field of application of the adaptive design methodology are clinical trials with several treatment arms, where promising treatments can be selected at an interim analysis. Using adaptive multiple test procedures the type I error rate can be controlled even if the selection rule, the number of selected treatments or the final sample sizes are not prefixed. Adaptive multiple testing procedures can also be used in adaptive designs with the option of population enrichment. In such designs a sub population may be selected in an interim analysis and further recruitment of patients is restricted to the selected subgroup. In the past few years adaptations proposed in regulatory applications may involve a hybrid or a complex form of various design features, such as reasonably likely surrogate or predictive biomarker, external control. This short course will share with some regulatory experiences in such adaptive designs in cardiovascular, renal, CNS and imaging drug trials.

Adaptive Clinical Trial Designs

- Group sequential designs
- Adaptive combination tests
- Multiple testing in adaptive designs
- Case Studies
- Some regulatory experiences

Short Couse	9:30AM - 12:45PM, August 26, 2019	Room #665, 5F

C. Accelerating Drug Discovery through Precision Medicine and Innovative Designs: Concepts, Rationale, and Case Studies

Instructor(s): **Sandeep M. Menon**, Pfizer Inc., Boston University, and Tufts University School of Medicine; **Weidong Zhang**, Pfizer Inc.

Precision medicine has paved the way for a new era of delivering tailored treatment options to patients according to their biological profiles. Advancement of the biotechnologies such as next generation sequencing technology (NGS) and other omics technologies have enabled us to interrogate a patient's many molecular biomarkers, and associate them with disease and drug responses. In addition, incorporation of biomarker information in the innovative clinical trial design has presented drug developers unprecedented opportunities to bring a successful drug to patients in needs.

The first part of this course will focus on the concept of precision medicine, biomarker discovery and its application in clinical trials. Comprehensive review of omics data and major technologies will be presented. Statistical considerations and challenges such as data normalization, dimension reduction and biomarker threshold development and using biomarker for decision making in clinical development will be discussed in details.

The second part of this course will focus on the strategy of the study design that is important to critically determine biomarker performance, reliability and eventually regulatory acceptance. A general overview of the concept and statistical methodologies and designs related to precision medicine will be presented. Specifically, we will discuss various designs including adaptive designs available at our disposal and its merits and limitations.

Short Couse	1:45 - 5:00PM, August 26, 2019	Room #663, 5F

D. Novel Adaptive Clinical Trial Designs for Immunotherapy and Modern Drug Development

Instructor(s): **Cong Chen**, Merck Sharp & Dohme Corp.; **Guosheng Yin**, University of Hong Kong; **Ying Yuan**, University of Texas MD Anderson Cancer Center

Following the success of PD-1 (or PD-L1) inhibitors, a flood of next generation immunotherapies with different mechanisms of action are being developed. While the expectation is high for these new immunotherapies, it is unrealistic to expect all of them to have the same success as their predecessors, especially given the improved standard-of-care. Innovative adaptive clinical trial designs provide a cost-effective and flexible way to improve the success rate of drug development. In this short course, we will present novel Bayesian designs for phase I and II clinical trials (including both single-agent and drug-combination trials), statistical strategies on phase 1 efficacy screening, adaptive 2-in-1 design for seamless Phase 2/3, Phase 3 adaptive designs for population expansion. We will introduce the freely available software and illustrate the application of the designs using real-world examples. This short course is suitable for statisticians and clinicians from industry, regulatory agencies and academia. Students of this short course are expected to not only apply the new methods learned to their studies but also think out of box when facing unique situations.

Short Couse	1:45 - 5:00PM, August 26, 2019	Room #664, 5F

E. Artificial Intelligence for Medicine and Health **Cancelled**

Instructor(s): **Mark Chang**, Boston University

Artificial intelligence (AI) or machine learning (ML) has been used in drug discovery for many years under name of bioinformatics, such as sequencing, annotating genomes, analysis of gene and protein expression and regulation, linking the biological and disease network to the symptoms and adverse events, identifying structure-activity relationships in discovery and designing new drugable molecules. AI has also been used for the prediction of cancer susceptibility (risk assessment), cancer recurrence/local control, and cancer survival. In analysis of clinical trial data, predicted individual patient outcomes for precision medicine, similarity-based machine learning (SBML) has recently been used in clinical trials for oncology and rare disease without the requirement of big data as most ML methods do. The introductory course will cover supervised, unsupervised, semi-supervised, and reinforcement learning methods, and swarm and evolutionary intelligences. It aims at conceptual clarity and mathematical simplicity. Provide R code for some of the supervised learning methods and discussion case studies.

Table of Contents:

(1) Overview AI methods in Medicine and Health

(2) Supervised Learning Method
Tree-Based Methods
Support Vector Machine
Artificial neural network for Deep Learning
- Perceptron
- Convolutional Neural Networks
- Recursive Neural Networks
- Long Short-Term Memory Networks
- Deep Belief Network

(3) Similarity Based Method:
Nearest-Neighbors Method
Kernel Method
Similarity-based machine learning
Implementation in R
Clinical Trial Examples

(4) Other AI methods and Future Perspectives
Unsupervised Learning
Reinforcement Learning,
Evolutionary Intelligence
Future perspectives

Goals: attendees will learn common AI methods in drug development and medicine, be able to use the AI methods with R for medicine, especially for clinical trials, and be able to interpret the results.

Short Couse	1:45 - 5:00PM, August 26, 2019	Room #665, 5F

F. Hot Topics in Clinical Trials: Multiple Outcomes and Benefit:risk

Instructor(s): **Scott R. Evans**, George Washington University; **Toshimitsu Hamasaki**, National Cerebral and Cardiovascular Center

We discuss two hot topics in clinical trials. In Part I we discuss the design and analysis of clinical trials with multiple outcomes. In Part II, we discuss benefit:risk evaluation in clinical trials by using outcomes to analyze patient rather than patients to analyze outcomes.

PART I: The effects of interventions are multidimensional. Use of more than one outcome offers an attractive design feature in clinical trials as they capture more complete characterization of the benefit and risk of an intervention and provide more informative intervention comparisons. The tutorial will focus on design and analysis of clinical trials with such multiple outcomes. The first part of the tutorial will focus on methods for clinical trial designs evaluating efficacy of two interventions with multiple primary endpoints, especially co-primary endpoints. “Co-primary” means that a trial is designed to evaluate if the test intervention is superior (or noninferior) to the control on all primary endpoints. We describe methods for power and sample size calculations in clinical trials with multiple endpoints including recently developed approaches. We include real clinical trial examples to illustrate the concepts and to help participants apply the methods in practice, and illustrate how to implement the methods using standard statistical software including R and SAS.

PART II: In the future, clinical trials will have an increased emphasis on pragmatism, providing a practical description of the effects of new treatments in realistic clinical settings. Accomplishing pragmatism requires better summaries of the totality of the evidence that allow for informed benefit:risk decision-making and in a way that clinical trials consumers—patients, physicians, insurers—find transparent. The current approach to the analysis of clinical trials is to analyze efficacy and safety separately and then combine these analyses into a benefit:risk assessment. Many assume that this will effectively describe the impact on patients. But this approach is suboptimal for evaluating the totality of effects on patients. In part II of the tutorial, we will describe a broad vision for the future of clinical trials consistent with increased pragmatism. Greater focus on using outcomes to analyze patients rather than patients to analyze outcomes particularly in late-phase/stage clinical trials is an important part of this vision. We discuss the desirability of outcome ranking (DOOR) and the partial credit strategy for design and analysis of clinical trials based on benefit:risk assessment. These strategies involve utilizing composite benefit:risk endpoints with a goal of understanding how to analyze one patient before trying to figure out how to analyze many. With a desire to measure and weigh outcomes that are most important from the patient’s perspective, we discuss using patients as a resource to inform analyses.

In-Conference Workshops

In-Conference Workshop	1:30 - 5:00PM, August 27, 2019	Room 509, 5F

I. Innovative Clinical Trial Designs for Personalized Medicine

Instructor(s): **James Wason**, Cambridge University and Newcastle University; **Florian Klinglmueller**, Austrian Medicines & Medical Devices Agency

The workshop will cover statistical methods for designing and implementing innovative trial designs for personalized medicine.

We will demonstrate adaptive multiple testing procedures defined by directed, weighted graphs that provide an intuitive visual tool for constructing multiple testing strategies that reflect the, often complex, contextual relations between hypotheses in clinical trials. These designs permit mid-trial design modifications based on unblinded interim data, while providing strong family wise error rate control in a wide range of scenarios including trials with multiple treatment comparisons, endpoints, subgroups, or combinations thereof. Examples of permitted adaptations are dropping of treatment arms, selection of subpopulations, and sample size reassessment. We will also discuss innovative approaches to the design and analysis of trials incorporating patient subgroups, including basket trials, umbrella trials and adaptive signature designs.

The practical implementation, operating characteristics, as well as regulatory considerations of such designs will be illustrated using real case studies.

In-Conference Workshop	1:30 - 5:00PM, August 28, 2019	Room 509, 5F

II. Adaptive Multi-Arm Multi-Stage (MAMS) designs in confirmatory clinical trials: a practical introduction to the statistical methodology and its application

Instructor(s): **Lingyun Liu**, Cytel; **Yannis Jemiai**, Cytel; **Hrishikesh Kulkarni**, Cytel

The development of new therapies has been challenging due to high cost and failure rate. More efficient trial design can save the time to market with less patients. Two arm group sequential design have been widely used for over 40 years which is more efficient compared to the traditional fixed sample design. Such design allows early stopping for overwhelming efficacy or futility, and therefore save sample size and development time. The natural generalization of two-arm group sequential designs is to compare multiple treatment arms to a common control in a multiple stages (MAMS designs). This course will discuss adaptive multi-arm multi-stage design in the confirmatory setting. We will start with the theory and design for two-arm group sequential designs. We will then extend the theoretical framework to MAMS design. We will discuss the various methods for controlling familywise error rate in face of adaptive treatment selection and sample size adaptation and the pros and cons for each approach. We will demonstrate how to design such studies with real clinical study and powerful software package. Practical considerations will also be discussed for successful implementation of such designed trials.

ISBS2019 Sponsors

Platinum Sponsors

JMDC PPD SNBL abbvie

Gold Sponsors

Takumi Information Technology NOVARTIS DOT WORLD

Silver Sponsors

日本メディア株式会社 Japan Media Corporation Cytel A2 Healthcare Eisai SAS

JMDC

Japan's Largest Claims Database.

☐ Payer-based DB

Data source : claim, enrollment information and health checkup from 130 + payers
Data period : January 2005 – the latest
Data volume : about 7 million enrollments

☐ Hospital-based DB

Data source : claim + DPC survey data from 190 + hospitals
Data period : April 2014 – the latest
Data volume : about 8 million patients

Our service

Data Provision service	Data Analysis service

JMDC Inc.

https://www.jmdc.co.jp/
pbmadmin@jmdc.co.jp
+81-3-5733-8614

Founded upon principles of excellence and innovation, PPD shares our clients' passion for delivering life-changing therapies to help people live healthier lives.

PPD SNBL

HELPING DELIVER LIFE CHANGING THERAPIES

St Luke's Tower 12F | 8-1 Akashi-cho, Chuo-ku
Tokyo 104-0044 Japan

www.ppdsnbl.co.jp

SOLVING THE WORLD'S TOUGHEST HEALTH CHALLENGES TAKES ALL OF US.

AbbVie starts with research and innovation to develop and deliver new treatments to manage some of the world's most difficult diseases.

Our global biopharmaceutical company builds our pipeline to provide solutions in therapeutic areas served by our proven expertise.

To make new solutions available to patients, we team with our scientific peers, physicians, governments and advocacy groups.

When we work together, the result is a remarkable impact on patients' lives and the healthcare systems which serve them.

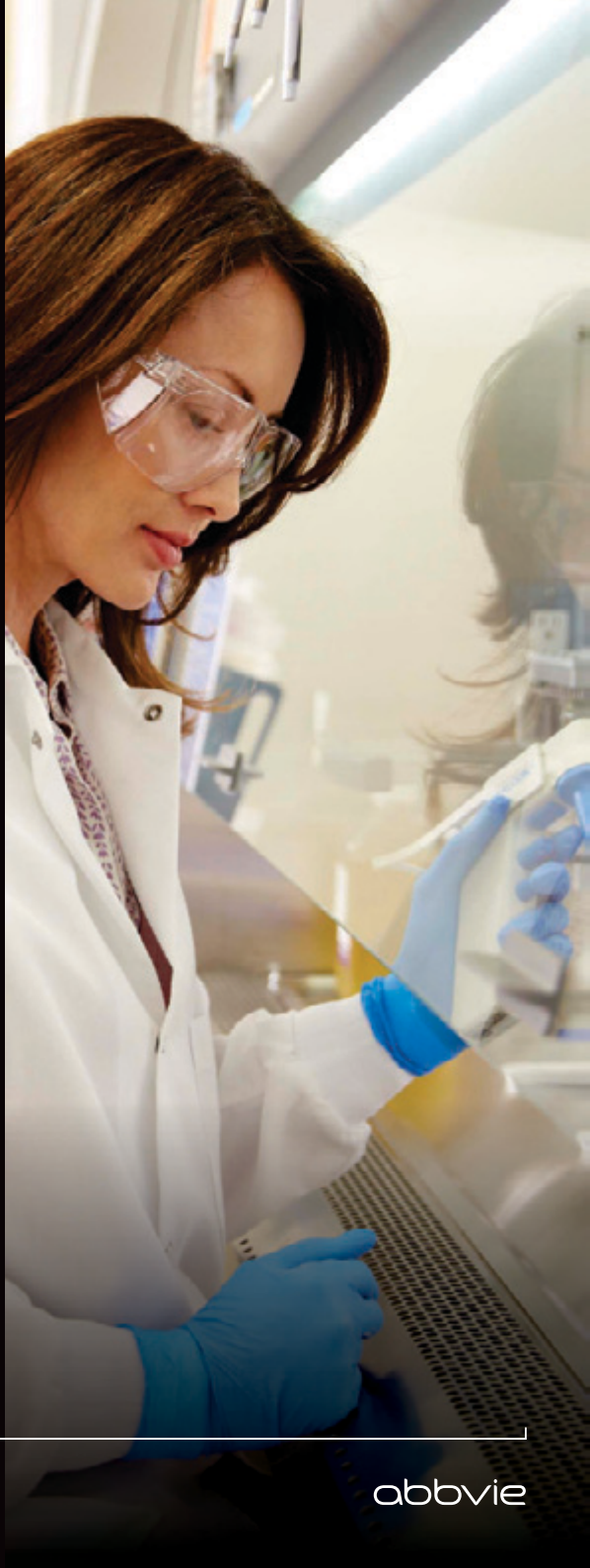
abbvie.com

170+ Countries

28,000+ Employees

21 R&D and Manufacturing Sites

ONE Priority: Our Patients



PEOPLE. PASSION.
POSSIBILITIES.

abbvie