BASS XXXI

Thirty-first Annual Biopharmaceutical Applied Statistics Symposium

November 4-6, 2024
Savannah, GA

Sunday, November 3, 2024

Registration 4:00–6:00 p.m.

Monday, November 4, 2024

Breakfast 8:00-9:00 a.m.
Tutorials 9:00–11:00 a.m.
Morning Break 11:00–11:15 a.m.
Tutorial 11:15 a.m.–12:00 p.m.
Lunch Break 12:00 p.m.–12:45 p.m
Tutorials 12:45–2:45 p.m.
Afternoon Break 2:45–3:00 p.m.
Panel Discussion 3:00–4:00 p.m.
Banquet Dinner 7:00–8:30 p.m

Tuesday, November 5, 2024

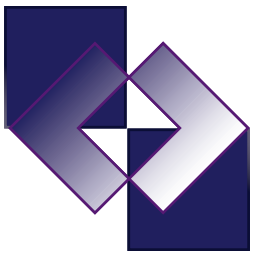
Breakfast 8:00-9:00 a.m.
Keynote Address 9:00-10:00 a.m.
Tutorial 10:00–11:00 a.m.
Morning Break 11:00–11:15 a.m.
Tutorial 11:15 a.m.–12:00 p.m.
Lunch Break 12:00–1:00 p.m.
Tutorials 1:00–3:00 p.m.
Afternoon Break 3:00–3:30 p.m.
Poster Session: 3:30–4:00 p.m.

Wednesday, November 6, 2024

Breakfast 7:30-8:30 a.m.
FDA Session 8:30-10:30 a.m.
Morning Break 10:30–11:00 a.m.
Tutorial 11:00 a.m.–12:00 p.m.
Lunch Break 12:00 p.m.-1:00 p.m.

Thursday, November 7, 2024

Virtual Short Course 1:00-4:00 p.m.
(Breaks planned with the instructor.)

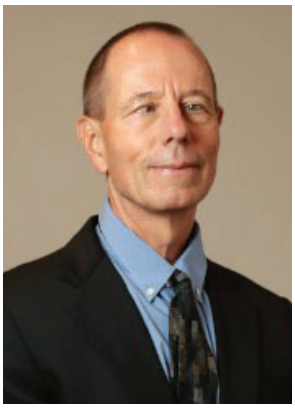


Keynote Address

Tuesday, 9:00 a.m.

Title: Building Your Career in Statistics: A Guide to Accomplishment, Happiness, and Longevity

Presenter: Dr. Craig Mallinckrodt, Pentara Corporation

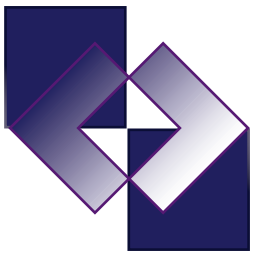


Craig Mallinckrodt is a Distinguished Biostatistician at Pentara, a CRO specializing in neurodegenerative research, where he contributes broadly to support research in Alzheimer's and other CNS diseases. Craig is a Fellow of the American Statistical Association and won the Royal Statistical Society's award for Outstanding Contribution to the Pharmaceutical Industry. Dr Mallinckrodt led several industry working groups, including the DIA Scientific Working Group on Estimands and Missing Data. He has authored four books, including books on longitudinal analyses, Estimands, Estimators, and Sensitivity Analyses and he has over 200 manuscripts on various statistical and clinical topics. Dr. Mallinckrodt led much of the early research into the MMRM analytic approach that resulted in this methods widespread use. He also has extensive experience in all phases of clinical development and led the statistical work that resulted in global regulatory approvals for drugs in 10 indications.

Abstract:

To have long, accomplished, and fulfilling careers we need to continue to grow and learn on the job. Although it can be hard to find the time, we have opportunities for continued statistical learning. However, the success and happiness we achieve in our careers is attributable to more than technical acumen. A variety of non-statistical aspects of our jobs are also key, and training opportunities in these areas is less common. The intent of this presentation is to outline a framework for career success and happiness built around continued learning and work-life balance that can be adapted to your own situation. The framework uses the analogy of a statistical career being like a car wherein statistical acumen is the engine that generates the raw power for our careers, with other features needed to translate that raw power into meaningful production. The presentation will include specific actions to become more productive, which is the

foundational skill for both work-life balance and to have time to become more creative, and to develop career plans and a continued learning strategy.



Monday Tutorials

9:00 a.m.

Title: A Case Study: The Adventure of the Missing Data

Presenter: Dr. Viral Panchal and Dr. Ryan Butterfield, Siemens Healthineers

Dr. Viral Panchal earned his DrPH in Biostatistics from the Jiann-Ping Hsu College of Public Health at Georgia Southern University, where he also completed his MPH. Prior to that, he received his Bachelor of Medicine and Surgery (MBBS) from Maharaja Sayajirao University, India. After completing his education, he completed his postdoctoral fellowship at the Medical College of Georgia. He then held the role of biostatistician at Duke University Medical Center before serving as an assistant professor and later as an adjunct professor of biostatistics at the University of North Carolina, Wilmington. Currently, he works as a staff biostatistician at Siemens Healthineers, where he leads and supports point-of-care projects in preclinical biostatistics.

Dr. Ryan Butterfield resides in St. Augustine, FL with family. On a personal note, he enjoys time with his family and golfing and reading as much as having little kids and his wife permits. His educational experiences resulted in a BS in Biomathematics and MPH in Biostatistics from Loma Linda University, an MBA from Jacksonville University, and a DrPH in Biostatistics from Georgia Southern University, where he was a BASS Fellow. Since graduating, he has worked as a Biostatistician in academia, hospital/clinical/University settings, varying levels of government (Local, State, Federal), and at several multi-international corporations at varying levels including 3M, Johnson & Johnson, and Edwards Lifesciences. He currently is a Senior Director of Clinical Biostatistics at Siemens Healthineers, where his team supports all stages of product development from R&D through Market entry.

Abstract:

This talk will be a case study on missing data, specifically missing data stemming from a longitudinal study conducted during the COVID pandemic. We will identify impacts on study execution, discuss ramifications on the data collection, and assess analytical strategies for potentially offsetting these issues. Discussion on potential impacts to study execution will focus on operational impact and responses as well as study design changes. We will examine ramifications of these responses on data collection and how potential systematic biases may be created in these scenarios. Analytical strategies to be discussed include techniques for assessing missing data, interpretation and components of the statistical justification for assessing missing data.

10:00 a.m.

Title: Up-Front Matching: A Recruitment Method for Prospective Observational Studies That Mimics Randomization

Presenter: Dr. Ibrahim Turkoz, Janssen Pharmaceuticals

Dr. Turkoz serves as the Scientific Director of Biostatistics at Johnson & Johnson Innovative Medicine Research and Development, LLC. Prior to his tenure at J&J, he gained extensive experience at various pharmaceutical consulting firms. With over 25 years of expertise in clinical research, Dr. Turkoz has contributed to all phases of clinical drug development. His current research interests include causal inference, comparative effectiveness, and cost-effectiveness, employing innovative methods such as pragmatic trials. As a founding member of the International Society for CNS Clinical Trials and Methodology (ISCTM), Dr. Turkoz has played a pivotal role in developing rating scales that assess patient functionality and clinical symptoms in CNS disorders. His work has also been instrumental in successful regulatory submissions across pain management, oncology, and device registries. Dr. Turkoz has published over 70 articles in peer-reviewed journals. He holds a Ph.D. in Statistics from Temple University.

Abstract:

As announced in the Federal Register notice published on October 2022, the US FDA is conducting an Advancing RWE Program, which seeks to improve the quality and acceptability of RWE-based approaches in support of new intended labeling claims, including approval of new indications of approved medical products or to satisfy post-approval study requirements. Data of high quality will be required to meet the high standards for regulatory actions based on relevance, integrity, and quality. Regulatory agencies demand transparency, traceability, and auditability of data. For these reasons, prospective observational studies (POS's) can generate real world evidence for regulatory purposes since their scientific validity is higher compared to other studies where there is no randomization. It is a major undertaking to control for selection bias and confounding in POS's. Corresponding analyses after collecting all the data points for selection bias are complex and these studies can also be expensive since not all patient level data are being utilized. For these reasons, we developed a recruitment method, called "up-front matching" , that will result in balance for selected baseline covariates that mimics what one would see with randomization. For POS's that are designed to compare treatments, this presentation will introduce up-front matching. Using the propensity score (PS) based on baseline covariates in a historical computer database, up-front matching (1) enrolls only patients whose PS is in the common support; and (2) performs frequency matching based on PS score strata. Up-front matching is compared to random sampling in a simulation. Across simulated studies, each with 200 patients per group, the distribution of the standardized mean difference for each of the covariates with up-front matching is well approximated by the normal distribution one would see with randomization (N(0, 2/200)), and with random sampling it is well approximated by N(std.diff, 2/200), where std.diff is the standardized difference in means. Use of up-front matching recruitment has the potential to (1) significantly enhance the scientific validity and statistical efficiency of observational studies, and (2) reduce the cost by following only patients in the common support.

11:15 a.m.

Title: Optimizing Drug Development Designs for Precision Medicine Considerations

Presenter: Dr. Gary Cline, Cline Consulting LLC

Gary has more than 25 years' experience in the pharmaceutical industry, including roles at ClinTrials, Procter & Gamble, ICON, Celgene and AstraZeneca where he has served on various governance committees. He has had various leadership responsibilities in strategic research and innovation, biostatistics, decision science, programming and data management across many therapeutic areas. He led Biometrics aspects for discovery and development programs, submissions and commercial activity for many products. Gary received his MS in Mathematics at Eastern Kentucky University and PhD in Statistics at the University of Kentucky.

Abstract:

Traditionally, early phase clinical trials seek to determine sufficient target engagement, proof of concept, and dose response. When considering precision medicine approaches, development plans need to determine the appropriate precision medicine group for a specific intervention which could be determined during any phase of study and/or utilize non-interventional data. The expected development plan approach is usually defined before the first human is dosed and primarily informed by preclinical and external data. The precision medicine group may be defined by categorical or continuous biomarkers or other nominal disease subgroups. Additionally, the consideration of a companion diagnostic is evaluated as well. We will explore various scenarios and options for drug development decision. One area to expand is response adaptive randomization to include response adaptive precision medicine population identification, i.e. Response Adaptive Randomization & Population (RARP). Ideally, we want to address the precision medicine scientific questions as early as possible within the drug development program; however, sometimes it may be advantageous to address the questions in Phase 3.

Lunch 12:00–12:45 p.m.

12:45 p.m.

Title: Real-World Data for Enhanced Drug Development in the Pharmaceutical Industry

Presenter: Mr. Greg Ginn, Realta Life Sciences

Greg is a distinguished Clinical Biostatistician with an extensive track record spanning over three decades across prominent sectors, including Medical Device, Pharmaceutical, Biotechnology, and Contract Research Organizations (CROs). His tenure in these industries has been marked by noteworthy achievements, where his expertise has made a substantial impact. In the realm of regulatory submissions, Greg has been a driving force behind the success of Premarket Approvals (PMAs), New Drug Applications (NDAs), Initial Public Offerings (IPOs), and pivotal Business Partnerships. His contributions have played an instrumental role in obtaining regulatory approvals, securing substantial investments, and forging vital industry collaborations. Furthermore, Greg has demonstrated remarkable leadership by overseeing Biostatistics and Clinical Data Management groups, instilling efficiency and precision in these critical functions. His initiatives have transformed several organizations by establishing and cultivating in-house capabilities for biostatistics, SAS programming, data management encompassing Electronic Data Capture (EDC), Interactive Data Capture (IDC), and Dynamic Data Capture (DDC), as well as medical writing. One of Greg's hallmark skills is his exceptional talent in talent management. He excels in recruiting and retaining top-tier staff, fostering an environment where excellence thrives. His ability to attract, develop, and nurture top talents is a testament to his capacity to build high-performing teams that drive results. In summary, Greg's outstanding career showcases not only a wealth of experience but also an unswerving commitment to innovation and excellence. His leadership, expertise, and vision have left an indelible mark on the biomedical and pharmaceutical industries.

Abstract:

In an era marked by rapid technological advances, evolving regulatory landscapes, and growing patient-centricity, the pharmaceutical industry is undergoing a profound transformation. Statisticians play a pivotal role in this evolving landscape by leveraging real-world data (RWD) to inform drug development, assess product safety, and enhance clinical trial efficiency. This presentation will explore the emerging methodologies, challenges, and opportunities in harnessing RWD for a data-driven approach to pharmaceutical decision-making.

1:45 p.m.

Title: Using Skeptical and Informative Priors in Bayesian Adaptive Clinical Trial Designs

Presenter: Dr. Roy Sabo, Virginia Commonwealth University

Dr. Sabo is a Professor in the Department of Biostatistics and is serving as Interim Chair in the Department of Epidemiology, and Interim Associate Dean for Faculty Affairs in the Virginia Commonwealth University School of Public Health. He is also the Director of the Biostatistics, Epidemiology and Research Design (BERD) Core in the C. Kenneth and Dianne Wright Center for Clinical and Translational Science. Dr. Sabo earned his Bachelor of Arts in Economics from Hamilton College (2000) and his Ph.D. in Statistics from Old Dominion University (2007). Dr. Sabo is a long-time collaborator with primary care researchers in the VCU Department of Family Medicine, focusing on cluster-randomized trials. He is also an expert in the design and analysis of population-level studies combining claims data with information from external sources. His methodological research focuses on Bayesian adaptive clinical trials, particularly with respect to using Decreasingly Informative Priors to avoid small-sample irregularities.

Abstract:

While adaptive study designs have the potential for improving clinical research, concerns remain over small-sample behaviors, particularly during early enrollment stages. Bayesian adaptive designs have the potential to remedy some of these concerns, but do not by themselves protect from small-sample irregularities. In this presentation we will illustrate the concept of decreasingly informative priors, designed expressly to suppress small-sample variability, while also allowing for gradual design adaptation. We first discuss the prior specification process, focusing on parameterizations that are simultaneously skeptical and informative, that gradually shift the weight of information from the prior toward the likelihood, all while maintaining a constant total effective sample size. We then show examples of how to apply these priors in adaptive clinical studies and trial designs, focusing on response-adaptive randomization, early termination, and dose-escalation.

3:00 p.m.

Panel Discussion: Biostatistics: Reflections of the Past and Projections for the Future

Panelists: Dr. Vipin Arora, Dr. Stephan Ogenstad, Dr. Robert Perera, Dr. Ryan Butterfield, and Dr. CV Daramaju

Moderator: Dr. Anthony Segreti, BASS

Banquet Dinner 7:00–8:30 p.m.

Tuesday Tutorials

10:00 a.m.

Title: Follow-up Planning for Time to Event Endpoints in Randomized Cancer Trials: Learnings for RWE Studies

Presenter: Dr. Shankar S. Srinivasan, Bayer Corporation

The speaker has a PhD in statistics from Virginia Tech and had prior majors in business and engineering. He has a total of more than 28 years of experience as a statistician in the pharmaceutical industry across all phases from phase II to phase IV, as well as non-clinical statistics. His recent exposure includes 18 years in medical affairs across multiple functions such as rheumatology, oncology and respiratory franchises with the last 14 years in Oncology. He has authored two patents and has published analytical methods as well as co-authored a number of journal publications and conference abstracts reporting pharmaceutical clinical trial data.

Abstract:

Trial planning details such as enrollment rates, follow-up for the last patient enrolled and accrual time could help in assessing adequacy of time-to-event analyses in real worldÂ studies by evaluating analogous concepts of cohort incidence rates, follow-up for last index patient and index period lengths. This tutorial will walk through the statistical design of event based oncology trials involving trade-offs between accrual periods, post-accrual monitoring periods and the study sample size with the intent of using these ideas in the real world context. This will help assess the adequacy of real world evidence in providing good estimates of effect where other elements of analytical rigor regarding missing data, biases and replicability have also been addressed.

11:15 a.m.

Title: Drug Target Discovery for Genetic Disorders Leveraging Multi-Omics Data and Comparability Assessment for Manufacturing Changes in CGT

Presenter: Dr. Lira Pi, PharmaLex GMBH

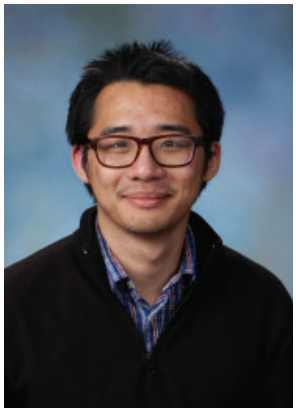
Dr. Lira Pi is an Associate Director in PharmaLex's Data Strategy and Quantitative Sciences Team. She obtained her Ph.D. in Statistics from The Ohio State University in 2013. Following her postdoctoral training at Duke University and University of Iowa, she joined PharmaLex in October 2020. Since then, she has applied her statistical knowledge across the entire spectrum of pharmaceutical R&D, including discovery, preclinical and clinical development, biomarker research, and translational medicine. Key areas of her research include integrating AI and Machine Learning into pharmaceutical R&D, constructing adaptive clinical design, and harnessing causal inference with real-world data and real-world evidence (RWD/RWE).

Abstract:

This presentation explores two statistical topics related to drug development for diseases with heritable risk factors. Understanding and translating molecular mechanisms behind genetic disorders early during development is critical for discovering potential drug targets. The most widely used method, genome-wide association studies (GWAS) to identify common single nucleotide polymorphisms (SNPs) are not sufficient to understand the complex molecular system. Multi-omics data enriched by chromatin accessibility, chromatin activity, and high-throughput chromosome conformation capture (Hi-C) data was further analyzed by the activity-by-contact (ABC) method to identify significant SNPs. The second topic discusses how to reliably assess comparability in autologous cell therapy manufacturing when there is a change in protocols. For testing comparability, FDA guidance suggests performing the two one-sided tests (TOST) with an equivalence margin defined before the study. Here, we propose to derive equivalence margins that ensure the "capability" , defined by the probability that the new process falls within limits of the old process.

Title: Challenges of Underdosing Control in Dose-Escalation Studies for DNA-Based Gene Therapies

Presenter: Dr. Quang Nguyen, Regenereon Pharmaceuticals Inc.



Quang Nguyen, PhD, is currently a Manager of Biostatistics at Regeneron Pharmaceuticals within the Scientific Insights group. He supports early phase clinical trials in the Genetic Medicines as well as exploratory data analysis projects across different therapeutic areas and indications. Previously, he received his Ph.D. from the Quantitative Biomedical Sciences program at Dartmouth College, focusing on statistical methods for understanding taxa-function relationships in human gut microbiomes.

Lunch 12:00–1:00 p.m.

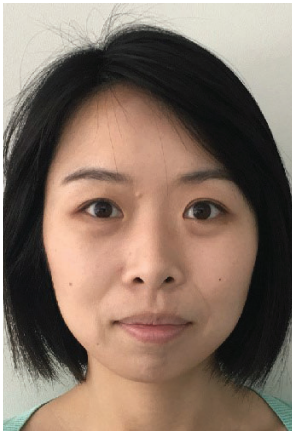
1:00 p.m.

Title: Gene Therapy in Rare Disease

Presenter: Chenkun Wang, Vertex Pharmaceuticals Inc.

Title: Assessing the Impact of Crossover on the Power for Overall Survival in Randomized Clinical Trials: A Simulation-Based Method

Presenter: Dr. Zhi Yang, Bristol-Myers Squibb Company



Dr. Zhi Yang is a Senior Manager in Biostatistics at the Cell Therapy Biostatistics Group of Bristol Myers Squibb. She has served as the protocol statistician and contributed to multiple cell therapy studies. Prior to joining BMS, Zhi worked as a Biostatistician at NanoString Technologies. Zhi earned her Ph.D. in Biostatistics from the University of Southern California in 2019 and her M.S. in Biostatistics from the University of Illinois at Chicago in 2015. Her research interests include dynamic borrowing from external control, adaptive trial design and Bayesian statistics.

Abstract:

Oncology clinical trials frequently allow patients in the control arm to cross over to receive the experimental therapy after disease progression (PD). Crossover may make a trial more appealing to patients, improving enrollment and retention rate. However, crossover can lead to change in overall survival (OS) afterwards therefore introduce uncertainty of the true treatment effect and power. It is important to assess crossover impact on OS power at the trial design stage. Given no existing methods to estimate the impact of crossover on OS power at the trial design stage, we propose a simulation-based method for the power estimation in a randomized controlled trial (RCT) with crossover at PD. By assuming exponentially distributed event times for death, PD, and drop-out, the power is estimated for a given sample size, hazard rates for death and PD, dropout rates, accrual rate, acceleration factor (AF), and follow-up period or observing a certain event number. AF describes the post-progression survival benefit due to treatment switching compared to the scenario had treatment switch not occurred. We illustrate statistical aspects and practical use of this method in designing a RCT in Cell Therapy to summarize the impact of readout time and AF on power, OS events, observed HR, median OS, and probabilities of observing a positive OS trend. We show that 1) the trends and magnitude of the crossover impact depend on the AF; 2) as the AF in a trial increases, the attenuation of the OS treatment effect and the impact on the power increases.

2:00 p.m.

Title: Statistical Techniques and Learnings from GVHD Treatment Trials

Presenter: Dr. Laura Morrison, The Emmes Company, LLC

Laura is a Senior Manager of Biostatistics for Emmes. She has worked on clinical trials for over 11 years, of which the last 8.5 years have been in cell and gene therapy. She has been the lead and statistical oversight of numerous trials and continues to do so. She has been involved in various FDA submissions and approvals including in the endocrinology, hepatology, and cell gene therapy areas, including a set of stem cell transplantation trials that went through the FDA and gained approval last year, and a few more than are set to go to the FDA in the next few months. In addition to her work at Emmes, she is also the Vice President of the Vancouver Island Life Sciences Association.

Abstract:

Hematologic malignancies are among the most common cancers and include lymphoma, myeloma, and leukemia. They are tumors caused by disruption of normal hematopoietic function. The first allogeneic stem cell transplant therapy to be approved by the FDA was for a nicotinamide modified allogenic hematopoietic progenitor cell therapy derived from a single umbilical cord blood unit for those with hematologic malignancies following myeloablative conditioning. In this presentation, we will review the patient population this product is for, why a product like this is needed, and the Phase I/II, Phase III, EAP, ISS, and ISE trials carried out with this product. The focus will be on the Phase III clinical trial that was the basis for the FDA approval. The Phase III trial used minimization as it's randomization scheme. We will review what minimization is and with what design types it can and cannot be used. We will discuss how randomness was added into the traditional minimization method, the advantages it had, as well as the challenges it presented the statisticians on the trial. An example of increased complexity resulting from the use of minimization is the requirement to use re-randomization which is computationally complex for the intent-to-treat analyses. We will investigate how re-randomization works, how it can be accomplished programmatically, and the difference between a standard testing confidence interval using

bootstrapping and a confidence interval based on re-randomization tests. We will look at the endpoints of the study and the associated statistical such as Cumulative Incidence with competing risks, Kaplan-Meier, log-rank test, and Gray's test. We will also discuss the primary endpoint test statistic which was to be a Mann-Whitney or Gehan-Wilcoxon depending on whether there were patients lost to follow-up by the time of the primary endpoint evaluation. For each of these, we will cover the formulas/methodology behind these methods and how to implement them in practice. Key results from the trial will also be presented.

Poster Session

3:30 p.m.

Poster Title: Use of Bayes Factors to Support Decision-Making in Early Phase Clinical Trials

Authors: Adlin Pinheiro, MA [1,2] ; Adrienne Lefeber, PhD [2] ; Lisa Hampson, PhD [2] , Abdelkader Seroutou, PhD [2] ; Astrid Jullion, PhD [2]

Affiliations:

[1] Boston University School of Public Health, Boston, MA

[2] Novartis Pharma AG

Abstract:

Introduction: In a phase II clinical trial with a dual-criterion design, a statistical significance criterion and a clinical relevance criterion must both be met to continue development of an experimental drug[1]. In the Bayesian framework, this design requires a high posterior probability that the treatment effect is greater than the null value and a moderate probability that the treatment effect is larger than a clinically relevant decision value. In the event of an inconclusive decision, occurring when only one criterion is met, other relevant factors must be considered to establish proof of concept. A measure that may help decision-making is the Bayes factor, which quantifies the evidence for two competing hypotheses[2].

Objective: In this study, we investigate the use of Bayes factors within the dual-criterion design. We assess what Bayes factors are expected

under various settings and evaluate decisions based on posterior probabilities and Bayes factors. We compare rates of misleading evidence and decision disagreement between decisions based on the two approaches.

Methods: Monte Carlo simulations were used to test for the difference between two population means under the null and alternative hypothesis settings for a fixed-n design, through Bayes factor design analysis[3]. The simulations involve drawing random samples of size n fixed and computing the Bayes factor using a Bayesian t-test, repeating this process 1000 times. We explore decision making when using a posterior probability threshold of 90% and various Bayes factor cutoff values.

Results: Under a point null hypothesis setting, the Bayes factor favors the null hypothesis for small effect sizes and small sample sizes and favors the alternative for mid-to-large effect sizes with large sample sizes when evaluating the statistical significance criterion. For small-to-large effect sizes, using a Bayes factor decision cutoff of 1 has lower rates of false negative results compared to the posterior probability approach when the sample size is small (<35 per arm). In addition, rates of disagreement are smallest when we use a Bayes factor cutoff of 1. Assessing the clinical relevance criterion results in smaller expected Bayes factors, and thus larger rates of misleading evidence.

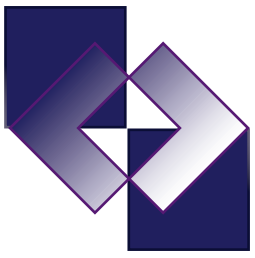
Conclusions: Bayes factors are not guaranteed to agree with conclusions based on posterior probabilities. Therefore, Bayes factors may provide additional information to inform decision-making in the scenario of an inconclusive decision within a dual-criterion design. However, careful consideration must be placed on choice of null and alternate hypotheses, Bayes factor decision intervals or cutoff values, and prior distribution selection as Bayes factors are more likely to favor the null hypothesis when it is false.

References:

1. Roychoudhury S, Scheuer N, Neuenschwander B. Beyond p-values: A phase II dual-criterion design with statistical significance and clinical relevance. Clinical Trials. 2018;15(5):452-461.

2. Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. Journal of the American Statistical Association, 90(430), 773–795.

3. Schönbrodt, F. D. & Stefan, A. M. (2019). BFDA: An R package for Bayes factor design analysis (version 0.5.0). Retrieved from https://github.com/nicebread/BFDA.



Poster Title: AI-Driven R Code Generation for Bayesian Simulations in Adaptive Clinical Trials

Authors: Haripria Ramesh Babu, MPH, J. Kyle Wathen, PhD, Boaz N. Adler, MPA

Keywords: AI-driven coding, Bayesian simulations, clinical trials, R code automation

Introduction/Goals: In adaptive clinical trials, Bayesian simulations are crucial for optimizing trial designs, but the complexity of R coding can be a significant barrier. This study introduces an AI-driven coding assistant developed to automate R code generation and integrate it with a cloud-based simulation platform. The assistant leverages advanced machine learning techniques to enhance coding efficiency and accuracy. The goal is to assess the assistant's effectiveness in facilitating Bayesian simulations, thereby enhancing trial design processes.

Methods: The AI-driven coding assistant was employed to generate R code for Bayesian simulations within adaptive clinical trials. The generated code was integrated into a cloud-based simulation platform, and key performance metrics—execution time, accuracy, and ease of integration— were compared with manually written scripts. The simulations focused on critical design elements, including treatment selection rules, posterior probability distributions, and decision criteria. Statistical analyses were conducted to validate the accuracy and efficiency of the AI- generated code.

Results: The use of the AI-driven assistant resulted in a 60% reduction in R code generation time while maintaining high accuracy and seamless integration with the simulation platform. Statistical analysis confirmed the significance of these improvements ($p < 0.05$). The AI-generated code demonstrated robust performance in Bayesian simulations, leading to improved trial design efficiency and more rapid optimization of trial parameters.

Conclusion: The AI-driven coding assistant offers a significant advancement in the automation of R code generation for Bayesian simulations in adaptive clinical trials. By improving the efficiency and accuracy of these simulations, this tool has the potential to enhance the overall trial design process. The technology's adaptability suggests broader applications in various clinical research contexts, providing a valuable resource for optimizing trial designs and potentially accelerating the development of new treatments. Additionally, this AI-driven approach opens opportunities for

collaboration and further research, enabling the continuous improvement of clinical trial methodologies.

Poster Title: Extending Integrated Multiple Adaptive Designs (IMADs) to Multiple Treatment Arms

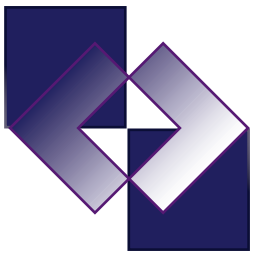
Presenter: Adam Funk

Advisor: Dr. Robert Perera

Affiliation: Virginia Commonwealth University, Department of Biostatistics, School of Public Health

Abstract:

Trial designs that improve efficiency in phase III randomized controlled trials (RCTs) are of interest to the Food and Drug Administration, pharmaceutical companies, and researchers alike. One class of designs that aims to improve the efficiency of RCTs while accounting for ethical considerations is integrated multiple adaptive designs (IMADs). IMADs incorporate both ethical allocation and efficiency through a multiple objective function which simultaneously optimizes response-adaptive randomization and sample size minimization. IMADs have been proposed for continuous and binary trials with two treatment arms. However, trials with more than two arms have some advantages as they increase the number of treatments that can be tested while often reducing overall costs. IMADs can contribute to this growing field by extending them to incorporate more than two arms. For the continuous case, we propose a natural extension of the multiple objective function to incorporate more than two arms. However, extensions to trials with more than two arms provide additional challenges such as a non-conforming response-adaptive randomization mechanism. Further, additional considerations are needed for trials with more than two arms, such as powering for an omnibus test, which tests for at least one difference between treatments, versus various sets of contrasts, which may compare a subset of treatment differences. The current proposed solutions, unresolved challenges, and future directions of our work are presented.



Poster Title: Local Time-to-Event Continual Reassessment Method for Drug Combination Studies

Authors: Li Liu, MS, Nolan A. Wages, PhD

Key words: Continual reassessment method: Late-onset toxicities; Likelihood-based design; Phase I trial; Dose optimization; Drug combination; Local model

Abstract:

Background: The primary goal of phase I cancer clinical trials is finding the maximum tolerated dose (MTD), defined as the dose with dose-limiting toxicity (DLT) probability closest to a pre-specified target (acceptable) DLT rate. Various methods have been proposed for both single-agent and drug-combination trials, including the continual reassessment method (CRM), Bayesian optimal interval (BOIN) design, and their extensions, such as the partial order CRM, the local CRM, and BOIN for drug combinations. These methods assume that DLTs are observed during a short-term window and are thus recorded before each new dose assignment. However, in modern oncology trials, patients often remain on therapy for longer periods and DLTs may be late-onset, occurring in later treatment cycles. Consequently, the observation window may be extended, while patient accrual remains rapid, leading to pending data at the time of new dose assignments. Although some methods have been developed for single agent trials with delayed toxicities, there is a lack of methods for drug combinations with delayed toxicities.

Aims: To propose a method to handle drug combinations with late-onset toxicities and evaluate its operating characteristics compared with similar methods.

Methods: The local CRM was extended to address late-onset toxicities. Adjacent doses of the current dose formed a local space and all possible toxicity orderings in this space were specified, with each ordering assigned a prior probability. This probability was sequentially updated at the time of each dose assignment. The time-to-event continual reassessment method (TITE-CRM) framework was used to model the DLT probability for each ordering, accounting for pending data from the extended DLT observation window. The DLT rate at each dose pair was updated using the weighted average of toxicity probabilities. At the time of each new dose assignment, the dose combination with an estimated DLT probability closest to the target rate that defined the MTD combination was selected.

Results: The local time-to-event continual reassessment method produced similar results to the local CRM in terms of percentage of correct selection (PCS) of the MTD, the number of patients treated at the MTD while substantially reducing the trial duration by adaptively utilizing pending data. It also demonstrated comparable performance to the partial order CRM in the presence of late-onset toxicities.

Conclusion: The proposed local TITE-CRM is suitable for phase I cancer drug combination trials with late-onset toxicities. It significantly decreases the overall trial duration while maintaining competitive operating characteristics for appropriate dose selection.

BASS Travel Stipend Award Winners:

- **Adlin Pinheiro - Boston University**
- **Corinne Mcgill - Medical University of South Carolina**
- **Lingling Wang - University of Alabama Birmingham**

Wednesday FDA/Industry Session

8:30 a.m.

Title: Applications of Unsupervised Learning for Plasma Concentration-Time Curves

Presenter: Dr. Junghi Kim, Division of Biometrics VIII, Office of Biostatistics, Center for Drug Evaluation and Research, Food and Drug Administration

Junghi Kim is a senior statistical reviewer at the Division of Biometrics VIII, Office of Biostatistics, in the FDA's Center for Drug Evaluation and Research. She provides statistical reviews for generic and biosimilar drugs and works on machine learning applications in drug development. She received her PhD at the University of Minnesota and received postdoctoral training at the University of Texas MD Anderson Cancer Center, before joining the FDA in 2018.

Abstract:

Pharmaceutical researchers are continually searching for techniques to improve both drug development processes and patient outcomes. An area of recent interest is the potential for machine learning (ML) applications within pharmacology. One such application that has not yet been given a close look is the unsupervised clustering of plasma concentration-time curves hereafter, pharmacokinetic (PK) curves. In this talk, we present our findings on how to cluster PK curves for similarity using hierarchical clustering. Specifically, we find clustering to be effective at identifying similar-shaped PK curves and informative for understanding patterns within each cluster of PK curves. Among existing methods, however, there are no dissimilarity measures specifically tailored to the task of clustering PK curves. We propose a novel dissimilarity measure, pkDip, to be used in the unsupervised clustering of PK curves. As an illustration, we apply several methods in a case study with 250 PK curves, using data from a previous pharmacogenomic study.

9:30 a.m.

Title: Emerging Digital Technology Data from Continuous Glucose Monitoring in Anti-Diabetic Product Clinical Trials

Presenter: Dr. Yoonhee Kim, Division of Biometrics II, Office of Biostatistics, Center for Drug Evaluation and Research, Food and Drug Administration

Dr. Yoonhee Kim is currently a team leader of the Diabetes and Obesity statistical review team in the Division of Biometrics II of the Office of Biostatistics, at the Center for Drug Evaluation and Research at US Food and Drug Administration. Dr. Kim has over 16 years of industry and government experiences for research and regulatory review works with applied statistical methods of high dimensional data analysis, random forest, and missing data analysis. Over the many years at FDA, Dr. Kim made substantial contributions for the Agency with her scientific and regulatory expertise focused in the therapeutic areas of diabetes. Her key areas of interest include continuous glucose monitoring data analysis, Bayesian borrowing analysis, estimands and clinical outcome assessment for evaluating the efficacy of pre-market anti-diabetic and weight management drugs.

Abstract:

Continuous glucose monitoring (CGM) involves using a wearable medical device to measure a patient's real-time blood glucose levels based on digital health technology (DHT). Because CGM produces glucose reading data every few minutes to track glucose values over multiple weeks at a time, CGM data is large scale and can be synthesized in multilayered structures from raw readings to CGM-derived endpoints. This abundance of CGM data offers opportunities to statisticians in regulatory agencies to analyze data at a more granular level and advance the data science for providing synthesized additional information for patients with diabetes. In addition, given the wide use of CGM in clinical practice, CGM-derived endpoints from clinical trials can be used to support efficacy and safety evaluation in the development of anti-diabetic products. To support the efficacy and safety of anti-diabetic products using CGM-derived endpoints, thorough collection of multi-week data, reliable quality assessment of ranging from raw readings to derived endpoints, and appropriate analytical methods

for regulatory decision-making are required. This talk will demonstrate a case study to contemplate these statistical aspects of CGM data based on DHT in drug development and a path forward for regulatory decision-making focused on handling missing data for CGM data quality assessment and subsequent analysis methods. This talk will provide more insights how the abundance of CGM data based on digital technology is playing a key role in clinical trials for data and regulatory science, and how the power of statistics functions in making valid CGM data and regulatory decisions for patients with diabetes.

11:00 a.m.

Title: Interpretable Machine Learning - Applications in Clinical Research

Presenter: Dr. Kao-Tai Tsai

Kao-Tai Tsai obtained his PhD in Mathematical Statistics from University of California, San Diego, and worked at AT&T Bell Laboratories to conduct statistical research, modeling, and exploratory data analysis. He later joined the US FDA and then various pharmaceutical companies, focusing on biostatistics, clinical trial research, and data analysis to address the unmet needs in human health. He has been active in the statistical profession and has engaged in numerous invited lectures, short courses, presentations, and seminars on practical statistical issues related to clinical trials. In addition, he was President of the NJ chapter of the ASA, served on the Board of Directors and various committees of ICSA and BASS, and organized symposia for several professional organizations. His recent research focuses on topics in clinical trials, biomarkers, big data analysis, and applications of statistical graphics in data analysis. He is the author of the book *Machine Learning for Knowledge Discovery with R, Methodologies for Modeling, Inference and Prediction*, published by Chapman and Hall/CRC in 2021.

Abstract:

The technologies in machine learning (ML) provide variety of important and useful tools in addition to the classical statistical data analysis methodologies. ML has received high degree of attention due to its ability to perform statistical estimation, modeling, and prediction to a wide variety of complex tasks. Â The deployment of ML systems in complex applications have led to a surge of interest in optimization not only for expected task performance but also other important criteria such as safety, unbiasedness, nondiscrimination, avoiding technical ambiguity or providing the right to explanation, etc. For ML systems to be used safely in various tasks or operational systems, satisfying these criteria is critical. However, these criteria often cannot be completely quantified, in such cases, a popular fallback is the criterion of interpretability: if the system can explain its reasoning, we then can verify whether that reasoning is sound with respect to these criteria. Generally, many of the ML methods require complicated computing and are not transparent in computations or background theory. Therefore, properly explaining the results they created becomes critical for it to be trusted and adopted by users. In the following, we describe various suggestions from researchers on the definition of explanation/interpretation of ML functionality, especially on the aspect of modeling. We also focus the ML applications on clinical trial research and describe the "Life Cycle of Data Analysis" based on the spectrum of procedures of complete clinical trial research operations.

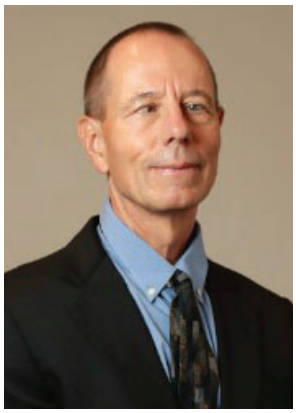
Box Lunch: 11:30 a.m.

Shortcourse

Thursday, 1:00 p.m. - 4:00 p.m.

Title: Building Your Career as a Statistician

Presenter: Dr. Craig Mallinckrodt, Pentara Corporation



Craig Mallinckrodt is a Distinguished Biostatistician at Pentara, a CRO specializing in neurodegenerative research, where he contributes broadly to support research in Alzheimer's and other CNS diseases. Craig is a Fellow of the American Statistical Association and won the Royal Statistical Society's award for Outstanding Contribution to the Pharmaceutical Industry. Dr Mallinckrodt led several industry working groups, including the DIA Scientific Working Group on Estimands and Missing Data. He has authored four books, including books on longitudinal analyses, Estimands, Estimators, and Sensitivity Analyses and he has over 200 manuscripts on various statistical and clinical topics. Dr. Mallinckrodt led much of the early research into the MMRM analytic approach that resulted in this methods widespread use. He also has extensive experience in all phases of clinical development and led the statistical work that resulted in global regulatory approvals for drugs in 10 indications.

Abstract:

This short course reviews and builds upon the ideas outlined in the Keynote address regarding the role of non-statistical skills in having a long, happy, and accomplished career through increased productivity, work-life balance, continued learning, and career development. We will take a more in depth look at specific things we can do to help ensure long-term success and happiness. We revisit the analogy of a statistical career being like a car, with our statistical acumen being the engine that generates the raw power for our careers, with other features needed to translate that raw power into meaningful production. Topics include improving: productivity and prioritization; innovation and creativity; critical thinking; and, influence and leadership. Each of these topics includes a section on the foundational principles and a section on putting those principles into practice. Connections between these individual skills are emphasized to see how the skills build upon each other leading to a whole that is greater than the sum of its parts. The

course will include real-life examples, with ample opportunity for questions. Participants will also be provided with the Author's book at no cost: Building Your Career As A Statistician: A Practical guide to a long, happy, and accomplished career.