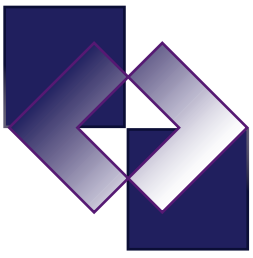
Thirty-second Annual Biopharmaceutical Applied Statistics Symposium

November 3-5, 2025
Savannah, GA



Schedule of Events

Sunday, November 2, 2025

Registration 4:00–6:00 p.m.

Monday, November 3, 2025

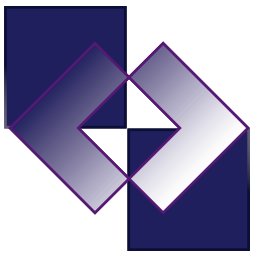
Breakfast 8:00–9:00 a.m.
Keynote Address 9:00–10:00 a.m.
Morning Break 10:00–10:15 a.m.
Tutorial 10:15–11:45 a.m.
Lunch Break 11:45–12:45 p.m
Tutorials 12:45–2:45 p.m.
Afternoon Break 2:45–3:00 p.m.
Tutorial 3:00–4:00 p.m.
Banquet Dinner 7:00–8:30 p.m

Tuesday, November 4, 2025

Breakfast 8:00–9:00 a.m.
Tutorial 9:00–10:45 a.m.
Morning Break 10:45–11:00 a.m.
Tutorial 11:00–12:00 p.m.
Lunch Break 12:00–1:00 p.m.
Tutorials 1:00–3:30 p.m.
Afternoon Break 3:30–4:00 p.m.
Poster Session: 4:00–5:00 p.m.

Wednesday, November 5, 2025

Breakfast 7:30–8:30 a.m.
FDA Session 8:30–10:30 a.m.
Morning Break 10:30–11:00 a.m.
Panel Discussion 11:00–12:00 p.m
Box Lunch & Departure 12:00–1:00 p.m.



Monday, November 3

Keynote Address

9:00 a.m.

Title: RCTS, Real-World Data and the Patent Wars: Lies, Damn Lies, and Biostatistics Case Study

Presenter: Dr. Stephan Ogenstad, Statogen Consulting

Dr. Stephan Ogenstad is an Expert Biostatistician at Statogen Consulting LLC, North Carolina since 2006, Founder and Principle of Innovalyst LLC, North Carolina, and In Silico Biosciences, Inc., MA. Dr. Ogenstad served as chief statistical advisor and reviewer to the Nobel Prize Committee for Medicine and Physiology and was before 2006, Vice President of Biometrics at Vertex Pharmaceuticals in Cambridge, Massachusetts, USA. He has 35+ years of experience with major pharmaceutical companies and distinguished university hospitals from around the world. Dr. Ogenstad earned a Ph.D. in Statistics from Stockholm University, Sweden, and served as professor of statistics at the Dept. of Statistics, Stockholm University. He is an Adjunct Faculty Member and Professor of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, a member of the International Advisory Committee of the University of North Carolina at Greensboro, USA, was President and Secretary of the North Carolina Chapter of the American Statistical Association, Program Chair for the Section on Statistical Consulting for the American Statistical Association, Member of the Clinton Health Access Initiative Clinical Advisory Board, Editor of MedCrave, Course Director of Biostatistics at Center for Professional Advancement, NJ, and President of the Swedish-American Chamber of Commerce of the Carolinas, Inc. Dr. Ogenstad today serves as a biostatistical expert witness in high-stakes pharmaceutical and medical device litigations, and is engaged in project collaborations with biotech, pharmaceutical, medical device companies, Massachusetts Institute of Technology, Georgia Southern University, and the Royal Institute of Technology, Stockholm, Sweden, where there are high-level strategy and problem-solving challenges, where his expertise can become fully utilized. He is Novo Nordisk's independent statistician for the LEADER and DEVOTE studies, in type-2 diabetes. Most of the project collaborations are as an expert biostatistician in pharmaceutical, device, and biotech submissions, where strategy and statistical specialtymethods, high-intensity modeling and simulation, machinelearning, risk analysis and decision-makingstrategies are critical.

Dr. Ogenstad also serves as chairman or independent statistician on Data and Safety Monitoring Boards. He interacts with FDA and EMA personnel at committee meetings and scientific advisory committee meetings on behalf of his clients. Dr. Ogenstad lectured and taught courses in stochastic processes, neural networks, survey sampling, and drug development and various other areas of statistics at Karolinska Institute and Swedish Academy of Pharmaceutical Sciences in Stockholm, Sweden. He is a frequent presenter at international conferences.

Dr. Ogenstad has had numerous professional affiliations as well as society memberships. He has been the president of the Swedish Society for Medical Statistics, council member of the Swedish Statistical Association, council member and co-founder of the European Federation of Statisticians in the Pharmaceutical Industry, and chair of the Massachusetts Biotechnology Council, USA, Biostatistics and Data Management.

During the past 25 years, he has held senior executive positions in nonclinical and clinical biostatistics and data management with AstraZeneca, Parexel, Amgen, and Vertex Pharmaceuticals.

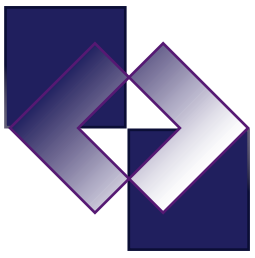
Dr. Ogenstad has been the lead statistician in the filing of 36 NDA, BLA, or PMA submissions in Europe, USA, and Japan. He has more than 100 publications, book chapters, and presentations in the areas of biostatistics and medicine, and is a reviewer of Statistics in Medicine, Communications in Statistics - Simulation and Computation, SAS Institute Inc., and Senior Editor for MOJ Clinical & Medical Case Report.

Abstract:

Randomized controlled trials (RCTs) are the gold standard—until they aren't. Retrospective studies are deeply flawed—except when they're not. Single-arm studies? Sometimes they're enough to get FDA approval. And then there's the creative world of real-world evidence, historical controls, and propensity score matching—where statistical gymnastics can make or break a study.

In this talk, I'll explore the messy reality behind study designs, from the rigid purity of RCTs to the Wild West of real-world data. Along the way, I'll ask some uncomfortable questions: Are we over-reliant on p-values? Are patents shaping the way we design studies? Is bias something we control, or just something we pretend to control?

With a mix of insight, controversy, and a bit of humor, this talk will challenge conventional wisdom and spark discussion about where biostatistics is headed in a world that's increasingly skeptical of the "one-size-fits-all" RCT model. Come for the statistics, stay for the war stories.



10:15 a.m.

Title: Leveraging Baseline Covariate Adjustment Methods in Oncology Time-to-Event Studies

Presenter: Dr. Daniel Backenroth, Johnson & Johnson

Daniel Backenroth is a fellow, biostatistics, in the Department of Statistics and Decision Sciences at Johnson & Johnson Innovative Medicine. He is based in Raritan, NJ. He has been at J&J since 2020, where he has focused on real-world evidence, causal inference and simulation-assisted trial design. He earned his PhD in Biostatistics from the Columbia University Mailman School of Public Health. Before J&J, he worked on the construction and analysis of clinicogenomic datasets at Flatiron Health, oncology data provider.

Abstract:

The primary endpoint in late-phase oncology trials is typically a time-to-event endpoint like overall survival or progression-free survival. Although covariate adjustment, in the form of stratification, is common in the analysis of these trials, much greater efficiency gains are now possible due to the development of the covariate-adjusted log-rank test (Ye et al. 2024, Lu and Tsiatis 2008). Using this test, together with data-driven selection of adjustment covariates, clinical trial sponsors can generate the same amount of evidence regarding the treatment effect of a novel investigational agent sooner, and with a smaller sample size, than is currently possible. Moreover, the test tests the same hypothesis as the unadjusted log-rank test, and comes with a hazard ratio estimator that has the same interpretation as the traditional Cox proportional hazards estimator, making it a "plug and play" solution to make time-to-event trials smaller and/or faster. In this talk we will discuss how machine learning techniques, together with publicly available trial data from Project Data Sphere and real-world data, can be used to generate and validate prognostic scores (e.g., linear combinations of baseline covariates) in a variety of oncology indications, which can be used as adjustment covariates for the log-rank test. We show how adjusting for these prognostic scores makes trials much more efficient than the traditional stratification approach. We also discuss how even greater efficiency gains are possible if multiple covariates are adjusted for. In this case, a machine learning technique called cross-fitting can be used to improve the finite sample properties of the covariate-adjusted log-rank test. We will show how to design realistic simulation studies that can be used to determine the expected finite sample properties of potential testing strategies in terms of both the Type I error rate and expected power gains.

Finally, we will discuss how the adjusted log-rank test can be used together with adaptive designs to take advantage of these efficiency gains and deliver high-quality evidence faster and with fewer patients.

11:00 a.m.

Title: Driving Operational Excellence: The Power of Statistical Innovation in Clinical Development

Presenters: Dr. Fei Chen, Johnson & Johnson

Fei Chen is a Senior Director of Biostatistics within the Statistical Decision Sciences (SDS) group, part of Global Development at Johnson & Johnson Innovative Medicine. Since joining Janssen in 2011, he has provided statistical support across all therapeutic areas at Johnson & Johnson. His work is centered on simulation-based clinical trial designs and optimizing clinical trial execution through statistical modeling and data-driven strategies.

Abstract:

This talk will examine how statistical innovation is transforming operational efficiency in clinical development. I will explore practical applications of advanced statistical methods in key areas such as patient recruitment, event projection, and clinical drug supply optimization. Through real-world case studies, the speaker will illustrate how close collaboration between statisticians and operational teams can drive cost savings, accelerate timelines, and enhance decision-making quality.

Attendees will gain actionable insights into the challenges and opportunities of embedding statistical thinking into trial operations, as well as strategies for fostering effective cross-functional partnerships. The session will conclude with a forward-looking discussion on the evolving role of statisticians in drug development and why embracing a collaborative, solution-oriented mindset is essential to maximizing impact in an increasingly complex clinical landscape.

Lunch 11:45 a.m.– 12:45 p.m.

12:45 p.m.

Title: Bile Acids Are the Strong Indicators for the Progression of Metabolic Dysfunction Associated Steatotic Liver Disease in Ethnically Diverse Populations

Presenter: Dr. Ekaterina Smirnova, Virginia Commonwealth University

Dr. Smirnova is an Associate Professor of Biostatistics. She received her PhD in Statistics in 2014 from the University of Texas at Dallas. She is an expect in the analysis of continuous signals from the functional MRI and wearable devices, and microbiome data. Her primary research focus is on the understanding the role of the microbiota involved in bile acid metabolism in Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD) and in the progression to Metabolic Dysfunction-Associated Steatohepatitis (MASH) and further to cirrhosis. Her long-term goal is to elucidate the role of microbiome in the development and progression of liver disease to facilitate the development of both diagnostics and therapeutics for the condition.

Abstract:

Metabolic dysfunction associated steatotic liver disease (MASLD), previously known as nonalcoholic fatty liver disease (NAFLD) is the leading cause of liver-related mortality worldwide. Despite the increasing metabolic dysfunction associated steatotic liver disease (MASLD) epidemic globally, current therapeutics for the condition have low response rates. This is mainly due to high heterogeneity of the disease, poor understanding of the factors contributing to the disease progression to metabolic dysfunction associated steatohepatitis (MASH), and the lack of understanding of biological mechanisms underlying MASLD progression. This is approached by investigating the mechanisms by which the gut microbiota generates different bile acid metabolites associated with disease progression from MASLD to MASH and to fibrosis. Intestinal and circulating bile acids play an important role in overall metabolic health, and have been linked to the development and progression of MASLD. The pool composition of circulating bile acids is largely controlled by levels and activities of bile acid 7α-dehydroxylating (7α-DeOH) gut bacteria and how they are regulated. We discuss the role of deoxycholate (DCA) bile acid and its derivatives in the development and progression of MASLD.

We further discuss the methodological and statistical challenges in (1) validating the previously observed relationship between fecal andcirculating DCA and metabolites and (2) linking those to the mechanisms of secondary bile acid formation in an ethnically diverse large-cohort study of MAFLD patients from USA, Turkey, Brazil and Singapore. This study will serve as a foundation for targeted bile acid-based biomarker studies and microbiome-based therapeutics for MASH.

1:45 p.m.

Title: Integrated Analysis of Pre-Marketing Safety Data, Statistical Considerations and Estimands *Part One*

Presenter: Dr. Katarina Hedman, AstraZeneca

Ph.D. Medicine, M./BS. Mathematics, Executive MBA. Statistical Science Director, AstraZeneca Gothenburg, Sweden. Over 20 years in life science industry including +10 years at AstraZeneca. Leader of the AstraZeneca statistical safety working group for the cardiovascular, renal and metabolic therapeutic area.

Abstract:

Integrated analyses across multiple clinical trials are crucial for comprehensive safety evaluations. The integration of studies offers significant benefits, such as enhanced precision in estimations. However, it demands a methodical approach to prevent the introduction of unnecessary biases in comparative analyses and to ensure that estimates remain meaningful. In this presentation, we employ the ICH estimand framework to guide the integration of clinical trials, discussing essential principles and statistical considerations vital for planning integrated safety analyses.

Break - 2:45 p.m.

3:00 p.m.

Title: Integrated Analysis of Pre-Marketing Safety Data, Statistical Considerations and Estimands *Part Two*

Presenter: Dr. George Kordzakhia, AstraZeneca

George Kordzakhia: Ph.D. Statistics, MS. Mathematics; over 15 years of experience at FDA, working in Center for Drug Evaluation and Research, leading statistical teams supporting multiple clinical divisions. From 2022 works at Astra Zeneca CVRM division as a Statistical Sciences Senior Director

Abstract:

Integrated analyses across multiple clinical trials are crucial for comprehensive safety evaluations. The integration of studies offers significant benefits, such as enhanced precision in estimations. However, it demands a methodical approach to prevent the introduction of unnecessary biases in comparative analyses and to ensure that estimates remain meaningful. In this presentation, we employ the ICH estimand framework to guide the integration of clinical trials, discussing essential principles and statistical considerations vital for planning integrated safety analyses.

Banquet Dinner 7:00 p.m. - 8:30 p.m.

Tuesday, November 4

9:00 a.m.

Title: Evaluating Frequentist and Bayesian Confidence Intervals for Study Size Adjusted Risk Difference

Presenter: Benjamin Duncan, AbbVie

Benjamin Duncan is a Director of Safety Statistics and Therapeutic Area Head Immunology Safety Statistics at Abbvie Inc. He has over 30 years of combined experience in the biopharmaceutical industry in safety statistics analysis as well as the design and analysis of clinical trials, across all phases of development in multiple therapeutic areas. At Abbvie, he works with and leads a team of dedicated safety statisticians who are responsible for partnering with Patient Safety in providing integrated safety and benefit risk analyses for regulatory submissions, periodic reporting, surveillance, medical safety assessments, and publications. Statistical research interests include safety signal detection and general drug safety analysis across various topics. He obtained his M.S. in Biostatistics from The University of North Carolina. He also holds a B.S. in Statistics from The University of Georgia.

Abstract:

The Study Size Adjusted (SSA) method is commonly requested by FDA as a method for combining studies for assessment of risk. Computation of the SSA risk difference is straightforward and is easy to implement. However, there are no established and agreed upon methods for the derivation of corresponding confidence intervals around the SSA risk difference. Therefore, a robust method that can handle large incidence proportions as well as small incidence proportions arising from sparse data and smaller sample sizes (including cases of 0 events) must be formulated. This presentation will focus on the evaluation of several proposed frequentist and Bayesian methods with regards to adequate coverage, low Type I error rate, and adequate statistical power to detect true differences.

10:00 a.m.

Title: Enhancing the DMC Data Package Using Open Source Software, Artificial Intelligence, and Large Language Models

Presenter: Dr. Melvin Munsaka, AbbVie

Melvin Munsaka is a Senior Director and Head of Safety Statistics at AbbVie in the Statistical Sciences Department. He obtained his PhD in Mathematical Statistics from Queen's University in Canada. He also holds an MSc in Mathematical Statistics and an MEd in Mathematics both from McGill University in Canada. Melvin has been in the pharmaceutical industry for more than 25 years. He co-leads some initiatives of the ASA BIOP Bayesian Scientific Working Group, ASA BIOP Safety Scientific Working Group, and the PHUSE Safety Analytics and Data Visualization and Open-Source Technology Sub-teams. He is the Publicity Chair of the Midwest Biopharmaceutical Statistics Workshop. He is Past Chair of the ASA Section for Statistical Programmers and Analysts and has been member-at-large of the Executive Committee of the ASA BIOP Section. Melvin is also a lecturer at the University of Chicago.

Abstract:

Data Monitoring Committees (DMCs) play a crucial role in ensuring participant safety by continuously evaluating adverse events and other safety signals during clinical trials. However, one of the significant challenges they face is managing the vast amount of data (tables, listings, and graphs (TLGs)) that must be reviewed. This overwhelming volume can lead to critical insights or trends being missed. Manual searches of clinical trial data are prone to errors and can hinder thorough analysis, impacting the DMC's ability to make timely decisions. To address this, DMC data packages should be purposefully designed to assemble relevant information that supports the decision-making process. There is a growing consensus that these packages need an overhaul in terms of outputs, organization, and structure. Leveraging modern open-source tools can facilitate faster generation of DMC packages and offer various delivery modalities for efficient and timely data insights. This includes incorporating functionality for interactivity, drill-down capabilities, and the use of emerging technologies such as artificial intelligence (AI) and large language models (LLMs). This presentation will focus on developing DMC data packages that utilize visual analytics, open-source tools, AI, and LLMs. Examples of TLGs and open-source tools that can integrate AI and leverage LLMs will be demonstrated.

Break 10:45 p.m.

11:00 a.m.

Title: A Structured Approach to Designing Pivotal Diagnostic Accuracy Studies for Noninvasive Tests Using Meta-Analytic Techniques

Presenters: Dr. Ryan Butterfield and Dr. Edward DeVol, Siemens Healthineers

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.

Though born and raised in central Ohio, Ed now lives in Dallas, Texas with his wife who is originally from the Middle East. They have four grown children who are scattered across three continents – Asia, Europe and North America. Ed attended Oberlin College and graduated with a BA in Mathematics and then went to The University of Michigan in Ann Arbor for his MSc and PhD in Biostatistics with a thesis on martingale methods for censored data. He also has an MBA from Duke University. Ed took a faculty position with the Mathematics Department at Oakland University in Rochester, Michigan following his PhD. He was then accepted a position at the King Faisal Specialist Hospital in Riyadh, Saudi Arabia. Following that he worked at the Baylor Healthcare System in Dallas as Vice President of Quantitative Sciences for almost ten years. Following his wife back to her home, he then returned to Riyadh and to the King Faisal Hospital for eight years. Since 2023, Ed has worked as a Principal Biostatistician with Siemens Healthineers working primarily to support evidence generation activities for the company.

Lunch 12:00 p.m. - 1:00 p.m.

1:00 p.m.

Title: Multiple Testing in the Context of Group Sequential Designs

Presenters: Dr. Yevgen Tymofyeyev and Dr. Michael Grayling, Johnson & Johnson

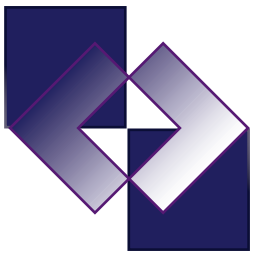
Yevgen Tymofyeyev is a Senior Scientific Director in the Statistical Modelling and Methodology group at Johnson & Johnson. In his current role, he serves as the statistical modelling lead for the Oncology Therapeutic area, implementing innovative designs and methods, including programs that utilize complex multi-stage designs with multiple hypothesis testing objectives, which have resulted in the successful submission of several clinical trials. He is actively involved in scientific collaborations in the field of randomization, adaptive design methodology, and software, which have led to an extensive list of publications, presentations, and implementation tools. Yevgen has 20 years of experience in pharmaceutical development applying complex methods in practice.

Michael Grayling is a Senior Principal Statistician within the Statistical Modelling and Methodology group at Johnson & Johnson, primarily supporting issues in Oncology and Immunology. Before joining J&J, Michael worked as a Research Fellow in Biostatistics at Newcastle University, where he developed a significant level of teaching experience and mentored multiple graduate students. His research interests include multi-arm multi-stage trials, crossover studies, and small sample sizes. He has published more than 50 papers in peer-reviewers journals and has authored a number of R packages and Stata modules related to trial design. Michael has a long record of teaching similar courses and delivering invited presentations focusing on adaptive design and multiple testing procures.

Abstract:

Multiple testing problems arise regularly in the design of clinical trials due to the presence of diverse sets of research hypotheses posed by multiple endpoints, treatment arms, subgroups, and combinations of these factors. Over the past decade, there has been a great expansion in the availability of methodology for performing sequential tests of multiple hypotheses. Amongst such methods, graphical testing in a group-sequential setting (see, e.g., Maurer and Bretz, 2013) has found particular utility, having now been leveraged in numerous studies. We will provide attendees with the necessary information to evaluate, select, and implement such a design in practice. This will include discussion of nuances related to planning the timing and triggering of interim analyses and comprehensive pragmatic detailing of analysis criteria.

Key options within this methodology will then be covered, including the utility of 'look back' analyses, how one can modify the alpha spending function for a hypothesis on updating the graph, and different alternatives for triggering interim analyses. We discuss both purely statistical and real-world considerations when selecting a design, and also detail how simulation can be used to estimate key marginal power quantities accounting for the correlation between all test statistics. We use recent trials as elucidating examples, to cover use cases with multiple arms, multiple endpoints, and multiple populations.



2:00 p.m.

Title: SMART-MC: Characterizing the Dynamics of Multiple Sclerosis Therapy Transitions Using a Covariate-Based Markov Model

Presenter: Dr. Priyam Das, Virginia Commonwealth University

Dr. Das pursued his bachelor's and master's degrees in Statistics at the Indian Statistical Institute, graduating in 2013. Subsequently, he embarked on a doctoral program in Statistics at North Carolina State University, where he concurrently completed a master's degree in Mathematics. He achieved dual degrees in 2016. Driven by a strong interest in Biostatistics, Dr. Das embarked on a postdoctoral fellowship at the University of Texas MD Anderson Cancer Center in 2017. Thereafter, his academic journey led him to Harvard Medical School, where he served as a research fellow during 2019-2022. Virginia Commonwealth University welcomed Dr. Das as an Assistant Professor in the Department of Biostatistics in 2022.

Dr. Das's current research interests include non-convex models in bioinformatics, scalable Bayesian methods in high-dimensional settings, Markov models for electronic health record (EHR) data modeling, dynamic treatment regimes, and graphical models in cancer proteomics. Among disease-specific modeling topics, Dr. Das is actively working on Multiple Sclerosis, Sickle Cell Disease, lung cancer, cancer prevention, Rheumatoid Arthritis, and Alzheimer's disease.

Abstract:

Treatment switching is a common occurrence in the management of Multiple Sclerosis (MS), where patients transition across various disease-modifying therapies (DMTs) due to heterogeneous treatment responses, differences in disease progression, patient characteristics, and therapy-associated adverse effects. To investigate how patient-level covariates influence the likelihood of treatment transitions among DMTs, we adopt a Markovian framework, Sparse Matrix Estimation with Covariate-Based Transitions in Markov Chain Modeling (SMART-MC), in which the transition probabilities are modeled as functions of these covariates. Modeling real-world treatment transitions under this framework presents several challenges, including ensuring parameter identifiability and handling sparse transitions without overfitting. To address identifiability, we constrain each transition-specific covariate coefficient vectors to have a fixed L2 norm. Furthermore, our method automatically estimates transition probabilities for sparsely observed transitions as constants and enforces zero transition probabilities for transitions that are empirically unobserved.

This approach mitigates the need for additional model complexity to handle sparsity while maintaining interpretability and efficiency. To optimize the multi-modal likelihood function, we develop a scalable, parallelized global optimization routine, which is validated through benchmark comparisons and supported by key theoretical properties. Our analysis uncovers meaningful patterns in DMT transitions, revealing variations across MS patient subgroups defined by age, race, and other clinical factors.

2:45 p.m.

Title: Harnessing Simulation to Inform Endpoint Selection in Cardiovascular Trials

Presenter: Dr. Stephan Ogenstad, Statogen Consulting - Co-Authored by Greg Ginn, ReAlta Sciences

Abstract:

The choice of primary endpoint in cardiovascular trials has profound implications for statistical power, interpretability, and regulatory acceptance. Traditional time-to-first-event analyses, such as the Andersen–Gill (AG) recurrent event model, offer established frameworks for capturing clinical outcomes. However, newer approaches like the Win Ratio, which prioritize hierarchical clinical importance across endpoints, may provide greater sensitivity in certain contexts. In this presentation, we describe a comprehensive simulation study designed to compare the operating characteristics of AG and Win Ratio methods under a range of plausible event scenarios typical of cardiovascular device and drug trials.

Our simulation framework incorporated recurrent event types (e.g., mortality, stroke, hospitalization) with varying frequencies, severities, and correlations, allowing evaluation of how treatment effects propagate across endpoints. We assessed bias, variance, type I error, and power for both methods, as well as interpretability and robustness to differing clinical event structures. Results highlight key tradeoffs: the AG model efficiently captures recurrent events and preserves statistical power when treatment effects are distributed across multiple event types, while the Win Ratio provides clinically intuitive comparisons when mortality or irreversible outcomes dominate.

These findings underscore the value of simulation as a decision-support tool in trial design, enabling sponsors and regulators to align statistical methods with clinical objectives. By demonstrating where AG or Win Ratio may be preferable, our work provides practical guidance for endpoint selection in cardiovascular trials, balancing methodological rigor with clinical relevance.

Break 3:30 p.m. - 4:00 p.m.

Poster Session 4:00 p.m. - 5:00 p.m.

Poster Title: Subgroup-Specific Overdose Control in Phase I Dose-Finding Studies

Authors: Rami Hawila and Dr. Nolan Wages
Virginia Commonwealth University, Department of Biostatistics

Abstract:

Background: Phase I trials aim to identify the maximum tolerated dose (MTD) while minimizing patient exposure to overly toxic doses. In the presence of patient heterogeneity, standard dose-finding methods may result in unsafe or suboptimal dosing, particularly in the early stages of a trial when data are limited.

Aims: This work proposes a subgroup-specific dose-finding design that integrates a toxicity-dependent feasibility bound into the continual reassessment method (CRM), with the goal of improving initial dose assignments.

Methods: The proposed CRM framework adopts a one-parameter power model to maintain parsimony and models subgroup differences using discrete shifts in dose recommendations. A dynamic feasibility bound is incorporated into the loss function to penalize overdosing more heavily early in the trial, gradually relaxing as additional participants are accrued and outcomes become available. Simulations were conducted across a range of clinically plausible toxicity scenarios to evaluate accuracy in MTD selection and dose allocation efficiency.

Results: The proposed design demonstrates favorable operating characteristics by effectively controlling overdoses early in the trial while accurately identifying subgroup-specific MTDs.

Conclusion: This work extends the shift-model CRM to incorporate toxicity-dependent escalation control for trials with heterogeneous patient groups. By combining model parsimony with dynamic overdose risk adjustment, the design enhances safety in subgroup-specific Phase I dose-finding.

Poster Title: Harmonic Fowlkes-Mallows Index for Medical Diagnostics Tests and Optimal Cut-off Point Selection of Binary Diseases

Authors: Parthkumar Rabari; Purbasha Biswas; Dr. Jacob Gakpo; Dr. Hani Samawi; Dr. Jing Kersey

Georgia Southern University, Department of Biostatistics, Epidemiology, and Environmental Health Sciences

Abstract:

Reliable differentiation between healthy and diseased populations is central to the evaluation of diagnostic tools in the biopharmaceutical sciences. This paper introduces the Harmonic Fowlkes–Mallows (HFM) index, a novel metric for quantifying diagnostic accuracy and guiding optimal threshold selection. The HFM index extends the classical Fowlkes–Mallows Index (FM) by incorporating a complementary Negative Fowlkes–Mallows Index (NFM), thereby integrating performance across both positive and negative classifications. These components are synthesized through a weighted harmonic mean, enabling balanced assessment of sensitivity and specificity. Unlike traditional criteria such as the F1-score or Youden's J statistic (Youden Index), the HFM index provides a unified framework that accounts for asymmetries in class relevance via a tunable parameter (β). Through simulation studies utilizing machine learning algorithms, the HFM index demonstrates strong performance in binary classification tasks and proves effective in selecting optimal decision thresholds. In this study, we utilized real-world breast cancer data to demonstrate the application of the HFM index alongside other established methods for evaluating diagnostic accuracy and determining optimal cut-off values for selected biomarkers.

Keywords: *Diagnostic test, Diagnostic accuracy, Biomarker(s), Cut-off point Selection, ROC, AUC, Youden Index, F-score, HFM, Machine learning, Breast cancer.*

Poster Title: Subgroup-Specific Overdose Control in Phase I Dose-Finding Studies

Authors: Moinul Ahsan and Dr. Nitai Mukhopadhyay
Virginia Commonwealth University, Department of Biostatistics

Abstract:

High-dimensional data often contain intricate nonlinear structures that challenge the assumptions of traditional parametric modeling techniques. In this work, we introduce an alternative prediction framework based on Topological Data Analysis (TDA) using the Mapper algorithm, designed to capture localized data geometry without prespecified global functional forms. Our approach constructs a topological graph with hierarchical clustering that segments the data via adaptive balanced cover. This enables predictions to be made by aggregating information from structurally similar regions of the data. We extend this framework to support both continuous and binary outcomes and propose a variable selection method based on permutation-based predictive loss, providing interpretable insights about the importance of covariates in terms of prediction. Our simulation study showed a better predictive performance compared to LOESS, GLM, and GLM-Q, but comparable performance with GLM-F. We further apply our method to a real-world clinical dataset from the Parkinson's Progression Markers Initiative (PPMI), identifying key clinical, imaging, and CSF biomarkers that distinguish Parkinson's patients with mild cognitive impairment from those with normal cognition at baseline. Our approach achieved comparable predictive performance (87.2\% accuracy; AUC = 0.797) to GLM (85.4\%, AUC = 0.796), Random Forests (85.4\%, AUC = 0.759), but better than LOESS (72.7\%, AUC = 0.736) while offering interpretable topological summaries of the disease phenotype. This work establishes Mapper-based TDA as a scalable, interpretable, and powerful alternative to black-box or traditional parametric models for prediction and feature selection in complex, high-dimensional data.

Poster Title: Statistical Methods for Improving Single-Cell RNA-seq Data Analysis

Authors: Alexandra Gerveni, Dr. Amy Olex, and Dr. Mikhail Dozmorov
Virginia Commonwealth University, Department of Biostatistics

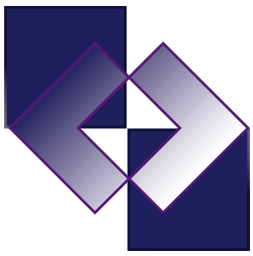
Abstract:

Single-cell RNA sequencing (scRNA-seq) technology allows to measure gene expression at a single cell level. Due to technological limitations, the data often contain problematic cells. These include cells with a low number of sequencing reads, doublets (high number of reads from multiple cells). Data from patient-derived xenograft (PDX) mouse models (mixture of human and mouse cells) contain another type of problematic cells, multiplets, with mixed reads from human and mouse. Current methods use fixed thresholds to detect such problematic cells. We developed a three-step statistical pipeline that classifies cells in a data-driven way.

Our method uses distribution-based thresholds across multiple quality dimensions. Step 1 identifies low number of reads cells using the lower tail of the distribution fitted to the total number of reads (mean - σ×SD). Step 2 detects potential doublets in the upper tail of this distribution (mean + σ×SD). Step 3 identifies multiplets as outliers in the distribution fitted to the percent of human/mouse reads. Our method has only one parameter, σ, that gives researchers control over stringency in detecting problematic cells, replacing arbitrary cutoffs with data- driven thresholds.

We tested our pipeline on a human-mouse mixing dataset from 10X Genomics containing 12,871 cells. The method successfully identified three types of problematic cells while maintaining all data with classification labels.

Our approach offers two main benefits: 1) data-driven thresholds making the method generalizable to various datasets, and 2) detecting up to three types of problematic cells as compared with methods targeting specific types. We plan to develop this method as an open-source R package to help researchers perform better quality control in their scRNA-seq analyses.



Poster Title: Spatially Co-expressed Gene identification through spatially varying networks

Authors: Ihsan Buker and Dr. Satwik Acharyya
The University of Alabama at Birmingham

Poster Preview:

SCG: Spatially Co-expressed Gene identification through spatially varying networks

Ihsan E. Buker[1] Satwik Acharyya[1]

[1] Department of Biostatistics, The University of Alabama at Birmingham, Birmingham, AL, USA

Introduction

- Spatial transcriptomics provides high-resolution, *in situ* maps of gene expression, revealing spatial architecture of tumor microenvironment.
- **Gap:** Most methods study spatially varying genes, deconvolution, and domain segmentation, but rarely focuses on spot/cell specific spatially varying gene co-expression networks.
- **Novel contribution:** We introduce **spatially co-expressed genes (SCG)** identification framework to study spatially varying marginal correlation, enabling discovery of spatial biomarkers which are potential targets for immunotherapy in cancer research.

Objectives

- **Build:** a flexible, scalable Bayesian spatial covariance regression framework.
- **Estimate:** spot/cell-specific gene–gene correlation across spatial domain of tissue.
- **Output:** Identification of Spatially Co-expressed Genes and spatially varying edges.

Spatial Transcriptomics (ST)

- Enables joint analysis of spatial architecture of tissue and biological process.
- **Per spot:** expression profile + spatial locations.
- **Resolution:** single cell (Visium HD, Xenium) and spot level (Visium).
- High dimensional spatial omics data with large n (cells), and p (genes).

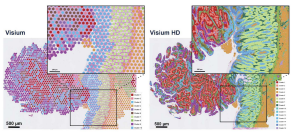
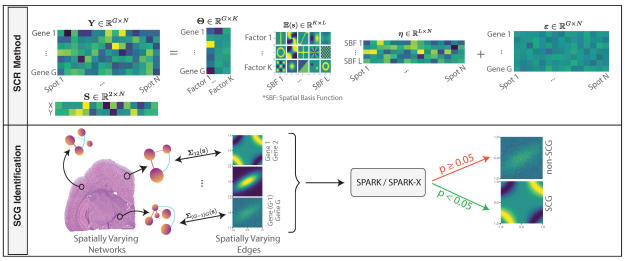


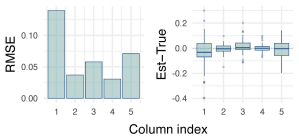
Figure: Resolutions of ST technologies [1]

SCR Method Overview

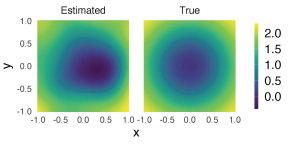


- **Model**
  - $\mathbf{y}_i = \boldsymbol{\Theta}\,\boldsymbol{\Xi}(\mathbf{s}_i)\,\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$
  - $\boldsymbol{\eta}_i \sim \mathcal{N}_K(\mathbf{0}, \mathbf{I})$
  - $\boldsymbol{\varepsilon}_i \sim \mathcal{N}_G(\mathbf{0}, \boldsymbol{\Sigma}_0)$; $\boldsymbol{\Sigma}_0 = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_G^2)$
- **Variational Inference**
  - Mean-field factorization
  - Coordinate-ascent variational inference for factor models [2].
  - Scales with high dimension.
- **Priors**
  - $\boldsymbol{\Theta} \sim$ Multiplicative Gamma Shrinkage Prior [3]
  - $\xi_{jk}(\cdot) \sim \mathcal{GP}(0, \kappa(\cdot,\cdot))$
  - $\sigma_g^{-2} \sim \mathrm{Ga}(a_0, b_0)$
- **Outputs**
  - $\boldsymbol{\Theta}$: global gene–basis structure; $\boldsymbol{\Xi}(\mathbf{s}_i)$: spatial trends
  - $\mathrm{Cov}(\mathbf{y}_i \mid \mathbf{s}_i) = \boldsymbol{\Sigma}(\mathbf{s}_i) = \boldsymbol{\Lambda}(\mathbf{s}_i)\,\boldsymbol{\Lambda}(\mathbf{s}_i)^\top + \boldsymbol{\Sigma}_0$
  - Apply SPARK [4, 5] on $\Sigma_{ij}(s)$ to identify SCG.

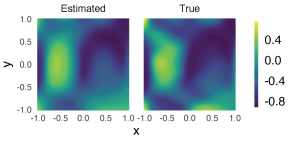
Simulation



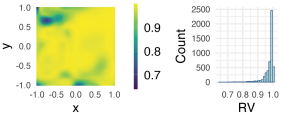
(a) Θ estimation by column: RMSE (left) and errors (right).



(b) Spatial basis $\Xi_{\ell=4,k=3}$: estimated vs. true.



(c) Spotwise corr. for gene pair (1,2): est. vs. true.



(d) RV: map (left) and distribution (right).

Figure: N=5,000 spots | G=50 genes | K=3 factors | L=5 spatial bases | SNR=10 | runtime ≈ 8 min.

Summary

- Proposed a scalable Bayesian framework for **spatial covariance regression (SCR)**.
- Estimation of spot/cell-specific spatially varying networks across spatial tissue domain.
- Identification of **SCG** to discover biomarkers.
- Study of spatially varying edges and disease propagation.

Future Directions

- Application to prostate cancer tissue to identify novel spatial biomarkers.
- Extension to multi-samples to estimate shared and sample-specific spatially varying networks.

References

[1] 10x Genomics Inc.
10x Genomics commercially launches Visium HD spatial gene expression assay.

[2] Massimiliano Russo Blake Hansen, Alejandra Avalos-Pacheco and Roberta De Vito.
Fast variational inference for bayesian factor analysis in single and multi-study settings.
Journal of Computational and Graphical Statistics, 0(0):1–42, 2024.

[3] Anirban Bhattacharya and David B Dunson.
Sparse Bayesian infinite factor models.
Biometrika, 98(2):291–306, 2011.

[4] Shiquan Sun, Jiaqiang Zhu, and Xiang Zhou.
Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies.
Nature Methods, 17(2):193–200, February 2020.

[5] Jiaqiang Zhu, Shiquan Sun, and Xiang Zhou.
SPARK-X: Non-parametric modeling enables scalable and robust detection of spatial expression patterns for large spatial transcriptomic studies.
22(1):184.

[6] Emily B. Fox and David B. Dunson.
Bayesian nonparametric covariance regression.
Journal of Machine Learning Research, 16(77):2501–2542, 2015.

Contact Information

- Email: iebuker@uab.edu

UAB
The University of Alabama at Birmingham.

BASS XXXII

Poster Title: A Comparison of Methods in Estimating Treatment Effect in the Presence of Rescue Medication in Randomized Clinical Trials

Authors: Ernestina Boateng, MS; Dr. Justin Leach; Dr. David Kimberlin; Dr. Gary Cutter; Dr. Inmaculada Aban

University of Alabama at Birmingham

Poster Preview:

A Comparison of Methods in Estimating Treatment Effect in the Presence of Rescue Medication in Randomized Clinical Trials

Ernestina O. Boateng, MS; Justin Leach, PhD; David W. Kimberlin, PhD;
Gary Cutter, PhD; Inmaculada Aban, PhD.
University of Alabama at Birmingham

Introduction

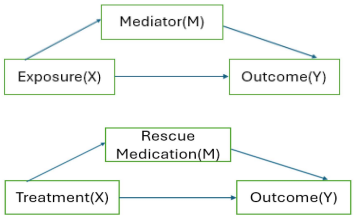
- Rescue medication (RM) refers to an adjunctive treatment administered to trial participants for alleviating worsening of symptoms in trials.
- Though it may benefit the patient, it may distort the true effect of the active drug.
- The International Council of Harmonization provides five strategies for handling intercurrent events where RM is a special case.
- Our focus is on the *hypothetical strategy* which envisions a scenario that RM did not occur.
- Using this approach, we can obtain the effect of the treatment on the outcome without any additional medications (direct effect).
- Most of the methods used to obtain the direct effect rely on asymptotic properties to estimate the treatment effect.
- However, there is limited information on how these methods perform in small samples.

Goal of the Study

To compare the performance of existing methods which estimate the direct effect of the treatment on the outcome at a fixed time point in small to moderate sample sizes.

Methods

- Approach is based on mediation analysis



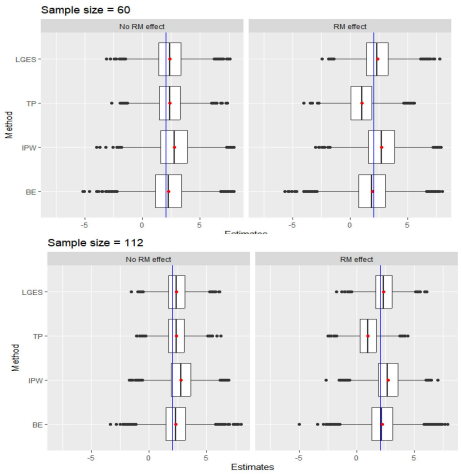
➢ Methods utilized includes:

- Loh G-estimation - relies on controlled direct effect.
- Balance Estimand - relies on natural direct effect.
- Inverse probability Weighting - Aims to balance the distribution of the covariates after discarding data for patients that used rescue medication.

Similarities between methods:

- All methods start with modeling the binary outcome of RM use.
- In an attempt to address the effect of the use of RM, the observed outcome values are adjusted accordingly.

Results



Conclusion

While all methods target different causal effects, Loh g-estimation performed consistently well across a range of scenarios based on bias and standard error of the estimate of treatment effects as well as simulated Type I error rates for testing treatment effects.

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

BASS XXXII

Poster Title: A Spline-Based Joint Model for Death and Progression with Mixed Censoring: Estimating PFS and CIF

Authors: Whitney Su and Dr. Yuan Wu
Duke University

Abstract:

Introduction: Progression-free survival (PFS) and cause-specific cumulative incidence functions (CIFs) are key endpoints in oncology and other disease areas where both progression and death are of interest. Standard nonparametric methods, such as the Aalen–Johansen estimator, require imputing partially observed progression times when monitoring occurs only at scheduled visits. This imputation can introduce bias and reduce efficiency, particularly when progression is interval-censored and death is right-censored. To address this challenge, we develop an imputation-free approach that directly estimates the joint distribution of progression and death while enforcing the clinical rule that progression must occur before death. This framework provides more accurate and clinically interpretable estimation of PFS and cause-specific CIFs.

Methods: We constructed a nonparametric maximum likelihood estimator (NPMLE) for the joint cumulative distribution function (CDF) of progression and death under mixed censoring. The joint model was expressed using I-splines/M-splines to enforce monotonicity and probability constraints and fitted via convex optimization in CVXR. From the estimated joint CDF, we derived PFS, CIF for death, and CIF for progression without imputation. Performance was evaluated through simulation studies under realistic visit-based interval censoring schemes with dependent event times generated from copula models. Comparisons were made to the Aalen–Johansen estimator with midpoint imputation.

Results: Simulation studies under copula-based data-generating models demonstrated that the spline-based estimator substantially reduced bias in estimating both the joint distribution function and the derived marginal distributions compared to standard imputation approaches. The method produced stable estimates of progression-free survival and cause-specific cumulative incidence functions across a range of dependence structures and censoring designs. In particular, the spline estimator yielded closer agreement with the true CIFs than midpoint-imputed Aalen–Johansen, which overstated progression risk. Additionally, the joint CDF estimator enabled direct evaluation of dependence between death and progression via Kendall's tau, showing accurate recovery of the underlying correlation structure.

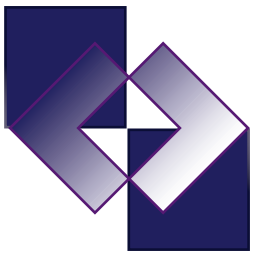
Conclusion: In simulations, the spline-based estimator showed markedly lower bias for progression-free survival and cause-specific cumulative incidence functions compared with midpoint-imputed Aalen–Johansen. The approach also accurately recovered dependence between death and progression, providing more clinically realistic estimates under interval censoring.

2025 BASS Scholars

- Moinul Ahsan - Virginia Commonwealth University
- Purbasha Biswas - Georgia Southern University
- Alexandra Gerveni - Virginia Commonwealth University

BASS Travel Stipend Award Winners:

- Ernestina Boateng - University of Alabama at Birmingham
- Ihsan Buker - University of Alabama at Birmingham
- Rami Hawila - Virginia Commonwealth University
- Whitney Su - Duke University



Wednesday, November 5

8:30 a.m

Title: Rare Disease Clinical Trials: Challenges and Opportunities

Presenter: Dr. Emily Morris, FDA Office of Biostatistics CDER

Dr. Emily Morris is a mathematical statistician in the Division of Biometrics IV in the Office of Biostatistics at the FDA's Center for Drug Evaluation and Research. She provides statistical support for the Division of Rare Diseases and Medical Genetics, which reviews drugs and biologics intended to treat inborn errors of metabolism. Prior to joining the FDA, Dr. Morris received her PhD in Biostatistics from the University of Michigan.

Abstract:

Rare diseases present unique challenges in drug development and statisticians play an important role in designing trials to overcome these challenges. One major challenge is the lack of information on natural history of diseases and the lack of early phase trials arising from having a limited patient population and feasibility constraints. Thus, when promising drugs are identified for rare diseases primarily based on non-clinical data, trials are often designed with limited information to inform decisions such as endpoint selection, sample size, and trial duration. The first portion of this talk will focus on ways that statisticians can help design a trial that has the highest chance of a conclusive result despite these limitations. The second portion of this talk will focus on the ways adaptive design may be used to address the lack of information at the design phase. Our research has investigated whether commonly used methods to control type one error in adaptive sample size trials may be applied to other types of adaptation to control type one error. Adaptations beyond sample size may be particularly useful in the rare disease context where an increase in sample size is considered infeasible. Instead adapting trial duration or other trial characteristics based on accrued information mid-trial may increase the likelihood of detecting a treatment effect.

9:30 a.m.

Title: Collecting and Applying Evidence to Interpret Clinical Outcome Assessments (COA)-based Endpoints in Clinical Trials

Presenter: Dr. Weimeng Wang, FDA Office of Biostatistics CDER

Weimeng Wang is a statistical reviewer and psychometrician in the Patient-Focused Statistical Scientists (PFSS) Group within the Division of Biometrics III, Office of Biostatistics (OB), Office of Translational Sciences (OTS), Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration (FDA). Dr. Wang received her PhD in Statistics, Measurement and Evaluation from the University of Maryland, College Park. Her current work at the FDA provides methodological expertise regarding the quantitative measurement aspects of patient experience data used in medical product development such as clinical outcome assessments.

Abstract:

Interpreting clinical outcome assessment (COA) scores is essential yet challenging in medical product development. When COA scores are used to construct endpoints in clinical trials, it is necessary to understand how the results of the COA-based endpoint correspond to a treatment benefit that is meaningful to patients. The U.S. Food and Drug Administration's Patient-Focused Drug Development (PFDD) draft Guidance 4 describes two families of methods: meaningful score differences (MSDs) and meaningful score regions (MSRs). The MSD approaches are a collection of methods which aim to translate differences on COA scores into differences in patients' experiences. The MSR approaches are a group of methods which seek to connect COA scores to corresponding patient experiences or judgements of experiences that are more easily interpreted than the native COA score. This presentation will introduce MSDs and MSRs and provide illustrative examples for each to aid stakeholders in applying and interpreting these approaches.

Break 10:30 a.m.

11:00 a.m.

Panel Discussion: *Artificial Intelligence and the Changing Role of the Biopharmaceutical Biostatistician*

Panelists:

Dr. Robert Perera
Dr. Ryan Butterfield
Dr. Vipin Arora
Dr. CV Damaraju
Dr. Stephan Ogenstad