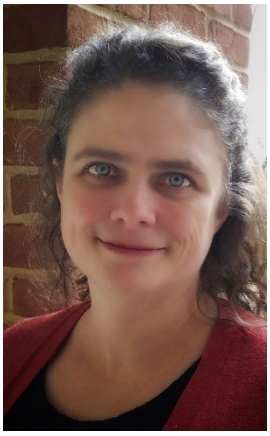
National Heart, Lung, and Blood Institute

NHLBI Biostatistics Workshop on Clinical Trial Designs: Innovative Endpoints and Futility Monitoring

Monday-Tuesday, September 18-19, 2023

Program Speakers

Session 1: Endpoint Choices in Clinical Trials



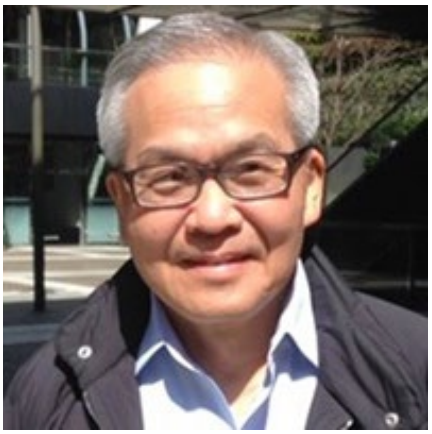
Dr. Laura Lee Johnson

U.S. Food and Drug Administration

Title: The Intersection of Patient Input and Statistical Methods to Inform Selection of Endpoints

Abstract: Considerations for selecting endpoints, whether used to assess benefit, harm, or tolerability, include the development of a well-justified rationale. We will primarily focus on endpoints constructed from fit-for-purpose assessments and discuss the elements needed to ensure an endpoint can support an inference of treatment effect within the context of a planned clinical trial. We will review examples including potential endpoint strategies when a disease affects multiple aspects of patient health and how the algorithm used to create an endpoint can elucidate uncertainty or render the summary of an assessment uninterpretable. We will focus on the importance of understanding how the endpoint corresponds to changes relevant to patients and decisionmakers, and consider how the endpoint, trial design, and analysis align with the clinical study objective to improve study planning and the interpretation of analyses.

Biography: Dr. Laura Lee Johnson is the director of the Division of Biometrics III in Office of Biostatistics at the U.S. Food and Drug Administration (FDA) Center for Drug Evaluation and Research (CDER). Her division supports the Office of New Drug's Office of Immunology and Inflammation and FDA efforts related to patient experience data, including Patient Focused Drug Development (PFDD). Her division works across FDA on methods to assess evidence including as part of COVID-19 efforts, rare disease initiatives, and clinical outcome assessment development. Prior to working at the FDA, Dr. Johnson spent over a decade at the U.S. National Institutes of Health (NIH) working on and overseeing clinical research and research support programs at the National Cancer Institute and what is now the National Center for Complementary and Integrative Health, including the CTSAs, NIH Collaboratory, and PROMIS. She co-directs the NIH Principles and Practice of Clinical Research annual course with over 10,000 participants in over 150 countries. Dr. Johnson earned her Ph.D. in Biostatistics from the University of Washington and is a Fellow of the American Statistical Association.



Dr. Lee-Jen Wei

Harvard T. H. Chan School of Public Health

Title: How to Choose an Estimand for a Comparative Study?

Abstract: The goal of conducting a clinical study is to obtain robust, clinically interpretable treatment effect estimate with respect to harm-benefit perspectives at the patient's level via efficient and reliable quantitative procedures. Estimand is a collective word which includes the choice of the study population, the definition of the endpoints, and the quantification of the treatment effect/difference in a comparative study. These are essential components for executing a successful clinical trial. In this talk, we will discuss various issues and concerns about the current practice via real-world examples. We then present simple remedies which would be appreciated by the patients, clinicians, and other stakeholders for treatment selection decisions. We will also discuss several open questions which need wisdom and input from the trialists and data scientists.

Biography: L. J. Wei is a professor of Biostatistics at Harvard University. Before joining Harvard, he was a professor at the University of Wisconsin, University of Michigan, and George Washington University. His main research interest is in clinical trial methodology, especially in design, monitoring, and analysis of studies. He has developed numerous novel statistical methods that are utilized often in practice. He received the prestigious Wald Medal in 2009 from the American Statistical Association for his contribution to clinical trial methodology. He is a fellow of American Statistical Associating and Institute of Mathematical Statistics. In 2014, to honor his mentorship, Harvard School of Public Health established a Wei-family scholarship to support students studying biostatistics. His recent research area is concentrated on translational statistics, personalized medicine under the risk-benefit paradigm via biomarkers, and revitalizing clinical trial methodology. He has more than 260 publications and served on numerous editorial and scientific advisory boards including data monitoring for governments and industry. L. J. Wei has extensive working experience in regulatory science for developing and evaluating new drugs/devices.

Dr. Scott Evans

Milken Institute School of Public Health, George Washington University

Title: The DOOR is Open: A Patient-Centric Approach to Clinical Trials

Abstract: Randomized clinical trials are the gold standard for evaluating the benefits and harms of interventions, though they often fail to provide the necessary evidence to inform medical decision-making. Primary reasons are: (1) failure to recognize the most important questions for treating patients in clinical practice, and (2) that traditional approaches do not directly address these most important questions, and subsequently using these recognitions as the motivation for the design, monitoring, analysis, and reporting of clinical trials. Standard approaches synthesizing information obtained from separate marginal analyses of each outcome fail to incorporate associations between or the cumulative nature of multiple outcomes in individual patients, suffer from competing risk complexities during interpretation of individual outcomes, fail to recognize important gradations of patient responses, and since efficacy and safety analyses are often conducted on different populations, benefit-risk generalizability is unclear. Cardiovascular event trials often utilize a time-to-first event primary endpoint where the events include e.g., death, stroke, and MI. Such endpoints are limiting, failing to recognize that fatal events are worse than nonfatal events, events with disabling sequelae are worse than events without disabling sequelae, and multiple events are worse than fewer events. The desirability of outcome ranking (DOOR) is a paradigm for the design, analysis, and interpretation of clinical trials based on a comprehensive patient-centric benefit-risk evaluation developed to address these limitations and advance clinical trial science. The DOOR uses outcomes to analyze patients rather than patients to analyze outcomes by comparing the experiences of trial participants in different treatment arms by the desirability of the patient-centric outcome. The DOOR paradigm, a recommended statistical analysis plan for research studies implementing DOOR and a freely available online tool for the recommended DOOR analyses, are described and illustrated.

Biography: Dr. Scott Evans is a Professor and Founding Chair of the Department of Biostatistics Bioinformatics and the Director of The Biostatistics Center at George Washington University. He is the: Director of the Statistical and Data Management Center for the Antibacterial Resistance Leadership Group (ARLG) funded by NIAID/NIH; the PI of the Coordinating Center for the Exercise and Nutrition Interventions to Improve Cancer Treatment-Related Outcomes (ENICTO) in Cancer Survivors Consortium funded by the NCI/NIH, and the co-PI of the Data Coordinating Center of the Clamp OR Delay among neonates with Congenital Heart Disease (CORD-CHD) clinical trial funded by the NHLBI/NIH. He is the Co-Chair of the Benefit-Risk Balance for Medicinal Products Working Group of the Council for International Organizations of Medical Sciences (CIOMS); Editor of a mini-Series on DSMBs for the NEJM Evidence; and the President-elect of the Society for Clinical Trials (SCT). He is a recipient of the Mosteller Statistician Award, the Zackin Distinguished Collaborative Statistician Award, the Founders Award from the American Statistical Association (ASA), an elected member of the International Statistical Institute (ISI), and is a Fellow of the ASA, SCT, and the Infectious Disease Society of America (IDSA).

Dr. Richard Cook

University of Waterloo

Title: Multistate Models as a Framework for Understanding Estimands and Assessing Intervention Effects

Abstract: Intensity-based multi-state models offer a useful framework for characterizing disease processes, the introduction of interventions, loss to follow-up, and other complications arising in the conduct of randomized trials of complex life history processes. Within the multi-state framework, the limiting values of common estimators of marginal process features (e.g., regression coefficients in cumulative incidence function regression models) can be evaluated to help in the interpretation of associated estimands. When intercurrent events can arise, we stress the need to carefully define the target estimand and the importance of avoiding targets of inference that are not interpretable in the real world. This not only has implications for analyses, but also the design of clinical trials where use of rescue therapy may be protocolized to aid in the interpretation of marginal features. Specification of utilities for different health outcomes, rescue interventions and other intercurrent events can also help synthesize information on complex processes to make simple treatment comparisons possible. This talk is based on joint work with Alexandra Buhler and Jerry Lawless.

Biography: Richard Cook is University Professor and Faculty of Mathematics Research Chair in the Department of Statistics and Actuarial Science at the University of Waterloo, and holds a cross-appointment in the School of Public Health Sciences. His research interests include the analysis of life history data, the design and analysis of clinical and epidemiological studies, and statistical methods for the analysis of incomplete data. He has published extensively in these areas and written two books with Jerry Lawless (*The Statistical Analysis of Recurrent Events,* Springer, 2007; *Multistate Models for the Analysis of Life History Data*, Taylor and Francis, 2018). He is a Fellow of the American Statistical Association. In 2018 he was awarded the Gold Medal of the Statistical Society of Canada, and in 2021 he was named a Fellow of the Royal Society of Canada.

Session 2: Endpoints in Cardiovascular Diseases

Dr. Phillip Yang

Stanford University School of Medicine

Title: Activity-based Prediction of the Clinical Outcome in Cardiovascular Patients Discharged from Hospital

Abstract: In our cardiovascular stem cell clinical trials, we have employed a broad category of standard endpoints: patient survey, cardiopulmonary exercise tolerance, cardiovascular function, and serum markers. However, there was never a remote, real-time collection of physical activity and biometric data of the trial subjects' daily routine at home. We hypothesized that this might provide a more accurate clinical assessment tool and prognosticate a patient's clinical outcome. Therefore, we employed a wearable, iPhone, AI interface, and back-end cloud server. We completed one clinical trial (ACT I, under review at JMIR) and are currently performing two more trials: ACT II and STOP-PASC (Long COVID, Pfizer Inc.). All the trials are conducted at Stanford University. The studies demonstrated novel trial methodology, digital wearable endpoints, and predictive AI algorithms for the trials.

Biography: Phillip C. Yang is a Professor of Medicine (Cardiovascular Medicine) at the Stanford University School of Medicine. He directs the Stanford Cardiovascular MRI Program and Cardiovascular Stem Cell Laboratory (Yang Lab). Dr. Yang received degrees from Stanford University and Yale University School of Medicine. Dr. Yang is a physician-scientist whose research focuses on innovation in myocardial restoration. His laboratory combines novel pluripotent stem cell biology and imaging modalities to advance cardiovascular therapeutics for heart failure. He leads multiple NIH, AHA, CIRM, and Stanford research grants, along with four clinical trials. He has received several prestigious awards, including the NIH Career Development Award, NIH Career Enhancement Award in Stem Cell Biology, NIH Mid-career Award, and multiple awards from both the American Heart Association and American College of Cardiology. He is a frequent guest speaker and session chair at national and international meetings.

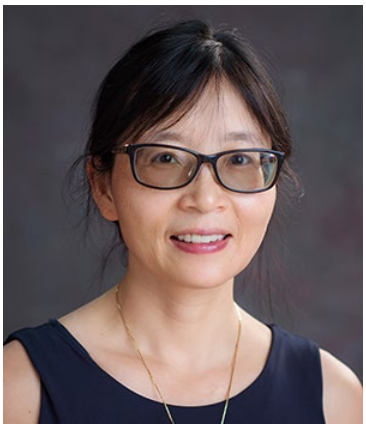
Dr. James Troendle

Office of Biostatistics Research, National Heart, Lung, and Blood Institute

Title: Use of Win Time for Ordered Composite Endpoints in Clinical Trials

Abstract: We describe several new variants of the Win Ratio that incorporate the time spent in each clinical state over the combined common follow-up. One version allows restriction so that death during follow-up is most important. Two other variants are described, one that creates a continuous outcome for each participant based on expected win times against a reference distribution and another that uses the estimated distributions of clinical state to compare the treatment arms. These new methods are designed to be closer to the overall treatment benefit/harm from a patient's perspective, compared to the ordinary Win Ratio. The new methods are compared to the composite event approach and the ordinary Win Ratio. When overall treatment benefit on death is substantial, the variants based on either the participants' expected win times against a reference distribution or estimated clinical state distributions have substantially higher power than either the pairwise comparison or composite event methods. The methods are illustrated by re-analysis of the trial Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training.

Biography: James F. Troendle is a Mathematical Statistician and Deputy Director of the Office of Biostatistics Research (OBR), Division of Intramural Research, NHLBI. Dr. Troendle received his Ph.D. in Mathematical Statistics from the University of Maryland (College Park) in 1991. Before joining NHLBI in 2010, he was at NICHD for 18 years. He is particularly interested in multiple comparison, longitudinal data analysis, and time-to-event analysis.



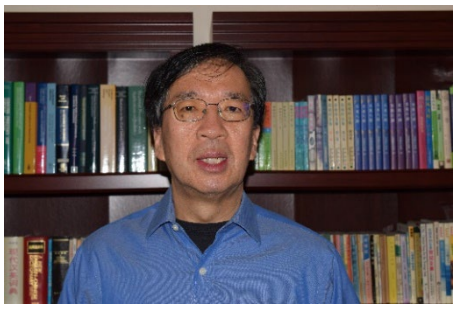
Dr. Tianxi Cai

Harvard T.H. Chan School of Public Health

Title: Toward Optimal Use of Surrogate Endpoints

Abstract: Motivated by increasing pressure for decision makers to shorten the time required to evaluate the efficacy of a treatment such that treatments deemed safe and effective can be made publicly available, there has been substantial recent interest in using an earlier or easier to measure surrogate marker, S, in place of the primary outcome, Y. To validate the utility of a surrogate marker in these settings, a commonly advocated measure is the proportion of treatment effect (PTE) on the primary outcome that is explained by the treatment effect on the surrogate marker. Model-based and model-free estimators for PTE have also been developed. While this measure is very intuitive, it does not directly address the important questions of how S can be used to make inference of the unavailable Y in the next phase clinical trials. In this talk, I will describe a framework for optimally using the information of surrogate S to maximally approximate the treatment effect on Y in a certain sense. I will illustrate our proposed procedures using an application to the Diabetes Prevention Program (DPP) clinical trial to evaluate the utility of hemoglobin A1c and fasting plasma glucose as surrogate markers for diabetes.

Biography: Tianxi Cai, Ph.D. (Biostatistics), is Professor of Biomedical Informatics (HMS) and Biostatistics (HSPH), and John Rock Professor of Population and Translational Data Science atHarvard University. Dr. Cai co-directs the VERITY Bioinformatics Core at Brigham and Women's Hospital and an Applied Bioinformatics Core at Veteran Health Administration. She is the founding director of the Translational Data Science Center for a Learning Health System at HMS and HSPH. She also directs the Big Data Analytics Core at Harvard Medical School, providing statistical and biomedical informatics support to both Harvard research community and external research groups, including VA and industry. Dr. Cai's research team has successfully developed statistical and informatics tools for analyzing complex big biomedical data from large scale studies including multi-institutional electronic health records, cohort studies, disease registries, genomic studies, and randomized clinical trials.



Dr. Song Yang

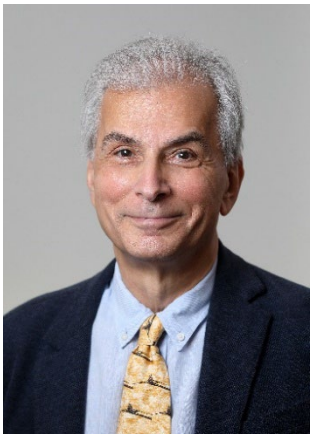
Office of Biostatistics Research, National Heart, Lung, and Blood Institute

Title: Testing and Estimation of Treatment Effects in Clinical Trials with Semi-Competing Risks Outcomes

Abstract: In clinical trials with time to event outcomes, often the outcomes of interest are a terminal event such as cardiovascular death, and one or more non-terminal events such as heart failure hospitalization. Traditionally, time-to-first-event analysis is used with the log-rank test or hazard ratio estimator derived under the Cox model for the first event. However, this approach is susceptible to bias and challenges in interpretation due to competing risks. It also uses data inefficiently. In recent years, more statistical models and procedures have been developed using effect measures such as the restricted mean survival time and its variants, the win ratio and its variants, and the Wei-Lachin test, as well as multivariate models such as the copula models and the illness-death models. In this talk, a brief review is provided on some of the approaches, and a few new inference procedures are also proposed. The data structure and testing and estimation problems are formulated to accommodate competing risks for both the terminal event and the non-terminal events. The presence of competing risks can otherwise lead to biased results and obscure interpretations.
The new methods can be rigorously justified, and they are illustrated using data from a few recent large cardiovascular trials. Recommendations are provided on how to choose an appropriate method in the design of a trial for practical treatment effect scenarios. Sample size, power, and interim analysis will also be addressed.

Biography: Song Yang is a senior mathematical statistician in the Office of Biostatistics Research, National Heart, Lung, and Blood Institute, NIH. Before joining NIH, Dr. Yang was a full professor and the Statistics Coordinator at Texas Tech University. His major research interests are survival analysis and clinical trials. In the last 30+ years, Dr. Yang has published numerous articles in methodology and clinical research in leading journals such as Ann. Statist., Biometrics, Biometrika, Clinical Trials, JASA, NEJM and Stat. in Med. He is also a frequent reviewer for these journals and for National Science Foundation proposals, and has co-authored the R packages YPmodel, ClinicalTrialSummary, YPInterimTesting, EventWinRatios, and YPmodelPhreg. Dr. Yang has been invited to present at numerous major conferences and university seminars. He is an associate editor for Lifetime Data Analysis and Stat. & Prob. Lett., a guest lecturer of the Foundation for Advanced Education in the Sciences at NIH, and a past guest co-editor for Stat. in Med. Since joining NIH, Dr. Yang has served as the program office statistician on more than a dozen large-scale clinical trials and studies and has received the NIH Director's Award and many other NIH awards for his service.

Session 3: Endpoints in Blood and Lung Diseases



Dr. Neal S. Young

Division of Intramural Research, National Heart, Lung, and Blood Institute

Title: Statistics: Practical Successes, Theoretical Problems: One Investigator's Perspective

Abstract: Our understanding and ability to effectively treat patients with bone marrow failure has been advanced by close collaboration between clinical and laboratory investigators – not usually known for their quantitative skills – and their statistician colleagues. For many decades, statistical considerations and innovations have been crucial to the design, implementation, and publication of many seminal studies in the Hematology Branch: multicenter trials sequential single center randomized controlled, single arm, and progressive protocols testing biologics and pharmaceuticals of aplastic anemia and related syndromes; population-based epidemiologic surveys in southeast Asia; machine learning applied to clonal hematopoiesis and in the differential diagnosis of constitutional versus acquired diseases; and in animal modeling and in vitro experiments by more basic scientists. As a result, the outcomes of previously fatal diseases have been reversed and understanding of their pathophysiology has deepened. Interesting and important issues arise from these collaborations. They include considerations of the ethics of clinical research in practice; the value of RCT; consequences of a highly bureaucratized, often cumbersome, and occasionally hostile regulatory environment; problematic reliance on animal, tissue culture, and molecular data for "translation" to clinical research; and the meaning and meaningfulness of a statistic attached to a practitioner confronting a patient – and the patient.

Biography: Neal Young is Chief of the Hematology Branch of the National Heart, Lung, and Blood Institute. His research interests are normal and aberrant hematopoiesis, autoimmunity in hematology, the genetics and genomics of aplastic anemia and related syndromes, and viral infections of blood forming cells. His career has been wide ranging, from basic molecular biology, virology, immunology, and cell biology to translational research, epidemiology, and pioneering interventional clinical trials. The Hematology Branch clinic is the major American referral center for marrow failure syndromes. Results from his work have deeply informed our understanding of the pathophysiology of human disease and development of effective treatments for aplastic anemia, paroxysmal nocturnal hemoglobinuria, myelodysplastic syndromes, and related diseases. He has published more than 450 original research articles and hundreds of reviews and chapters, including many monographs and more than two dozen articles in the New England Journal of Medicine. His trainees are the current leaders and international experts in the field of marrow failure. His accomplishments have been recognized by the American Society of Hematology with the E. Donnall Thomas and Beutler Prizes, and awards such as the Adolfo Storti Award, the Erasmus Prize, and lifetime honorary membership in the Mexican Society for Hematology. For public service, he received the Heyman Service to America Award for civil service and the Vietnam People's Award for his innovative teaching program in that country. Dr. Young recently completed a Visiting Fellow at New College in Oxford.

Dr. Anna Bellach

Office of Biostatistics Research, National Heart, Lung, and Blood Institute

Title: Regression Analyses for Competing Risks Endpoints in Lung and Blood Diseases

Abstract: Graft-versus-host disease (GVHD) is a major complication of a bone marrow transplant as a treatment for leukemia. On the one hand, GVHD can decrease the risk of relapse when white blood cells from the donor attack the patient's cancer cells. On the other hand, it can increase the risk for death in remission. It is therefore important to take account for the competing risks aspect of the data, and to analyze the effect of treatment and other covariates on both endpoints separately.
A general semiparametric regression model for the subdistribution hazard of a competing risk can be derived from a weighted likelihood function. Applying the new method to a large bone marrow transplant dataset, we assess how different donor types affect recurrence of leukemia and death.
Additional challenges occur with recurrent events and competing terminal events. We discuss how the general approach for competing risks data can be extended to establish a semiparametric regression model for the marginal mean of a recurrent event. This approach is relevant, for example, in chronic obstructive pulmonary disease (COPD) trials. An example is the STATCOPE trial that investigated the effect of Simvastatin in the prevention of COPD exacerbations in the presence of death as a competing event. We discuss how the proposed method can be applied to improve the statistical analyses of such trials.

Biography: Anna Bellach is a Mathematical Statistician at the Office of Biostatistics Research (OBR), Division of Intramural Research, NHLBI. Dr. Bellach received her Ph.D. from the University of Copenhagen. Her Ph.D. program was funded by a prestigious Marie Curie scholarship from the European Commission. A part of her Ph.D. research Dr. Bellach conducted at the University of North Carolina at Chapel Hill as a visiting scholar. Before joining OBR, she received additional training as a postdoctoral researcher at the University of Washington and at the Fred Hutchinson Cancer Research Center. Dr. Bellach's research interests include competing risks, recurrent events, semiparametric models, and empirical process theory.

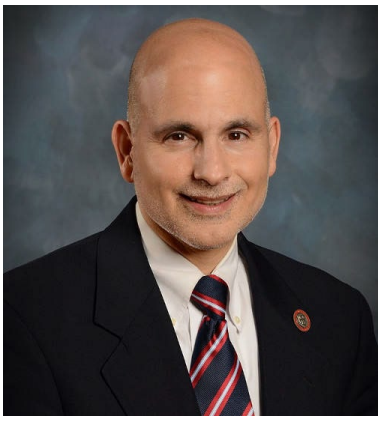
Dr. Ying Lu

Stanford University School of Medicine

Title: Using the Desirability of Outcome Ranking (DOOR) Approach to Construct Composite Endpoints from Multiple Outcome Domains

Abstract: Complex disorders affect multiple symptom domains measured by multiple outcomes. Successful treatments may affect one or several domains that may vary among patients. Composite endpoints that integrate outcomes across multiple domains help the evaluation of totality of treatment benefits. In this talk, we will present a general approach to construct a composite endpoint from multiple domains. This approach integrates multiple ordered outcome measures according to their relative ranking in an evaluation system. The ranking of outcome variables can be defined in a protocol (a shared decision making (SDM) trial) or vary by treatment approaches (such as for traditional Chinese medicine trials), or by patient preferences (such as the Patient-ranked Order of Function (PROOF) score for Amyotrophic Lateral Sclerosis (ALS) trials). Using the desirability of outcome ranking (DOOR) approach, we can construct the Mann-Whitney U-statistics to average all possible pairs of trial participants to estimate the probability of a treated participant having more desirable outcomes than a control participant. This approach has the advantage of flexibility in how many domains can be integrated, independent of measurement units, and improvement in relevance of efficacy and statistical power. We demonstrated this approach using the results from the ENHANCE-AF trial (NCT04096781), which evaluated a novel SDM pathway for patients considering anticoagulation for stroke prevention, and the follow-up data from the development of the PROOF in prediction of ALS patient survival. We will discuss challenges in using this approach and strategies to address them. The presentation is based on collaborations with Professors Lu Tian, Paul Wang, and Randall Stafford at Stanford University and Professors Ruben van Eijk and Leonard van den Berg at the University Medical Center, Utrecht, the Netherlands.

Biography: Ying Lu is Professor of Biomedical Data Science, and by courtesy, of Radiology and of Epidemiology and Population Health, Stanford University School of Medicine. He is a Co-Director of the Stanford Center for Innovative Study Design and the Biostatistics Core of the Stanford Cancer Institute, and co-Leader of the Biostatistics, Epidemiology and Research Design (BERD) resources of Stanford SPECTRUM. Before his current position, he was the director of Veteran Affairs Cooperative Studies Program Palo Alto Coordinating Center (2009-2016) and a Professor of Biostatistics and Radiology at the University of California, San Francisco (1994-2009). His research areas are biostatistics methodology and applications in clinical trials, statistical evaluation of medical diagnostic tests, and medical decision making. Dr. Lu is an elected Fellow of the American Association for the Advancement of Science (AAAS) and the American Statistical Association (ASA) and has more than 300 peer-reviewed journal publications.



Dr. Joseph Cappelleri

Pfizer Inc.

Title: Development and Psychometric Evaluation of a Novel Tool for Assessing Patient Perception and Preference for Haemophilia Treatment (HaemoPREF)

Abstract: *Introduction*: Prophylactic use of treatment is important for good outcomes in haemophilia, yet adherence can be sub-optimal. To better understand the relationship between treatment adherence and patients' beliefs about treatment, there is a need to quantify patients' treatment attitudes.

Aim: To develop a brief, clinically relevant, patient-reported outcome (PRO) to measure ease of use and patients' preference for haemophilia treatment.

Methods: A 40-item questionnaire was completed by male adults with haemophilia A from Austria, Germany, Italy, Spain, and the United Kingdom. Robust statistical methods for item evaluation including item-level statistics, dimensionality analyses, and input from clinical and outcomes experts were used to inform item reduction. Retained items were subjected to psychometric evaluation including exploratory factor analysis (EFA), known-groups validity, and internal consistency reliability.

Results: 273 patients completed the questionnaire. Of the 40 items, 28 items were flagged for possible deletion based on item-level statistics, three of which were retained due to clinical relevance. Two items had acceptable statistical performance but were deleted based on low clinical relevance. A total of 13 items were retained. EFA produced a conceptually defined 5-factor solution. The survey had acceptable known-groups validity and internal consistency. Refinements were made to wording and scoring, and one new item was added to assess general ease of use, resulting in a 14-item questionnaire – the HaemoPREF. An independent follow-up study with 89 patients added support for the validity and reliability of the measure.

Conclusions: Measurement properties of the HaemoPREF support the instrument to evaluate patient perception and preference for haemophilia treatment. Further psychometric evaluation is encouraged to confirm the measurement properties of the scale.

Biography: Joseph C. Cappelleri earned his M.S. in statistics from the City University of New York, Ph.D. in psychometrics from Cornell University, and M.P.H. in epidemiology from Harvard University. Dr. Cappelleri is an executive director of biostatistics at Pfizer Inc. As an adjunct professor, he has served on the faculties of Brown University, University of Connecticut, and Tufts Medical Center. He has delivered numerous conference presentations and has published extensively on clinical and methodological topics. Dr. Cappelleri is an elected Fellow of the American Statistical Association (ASA), elected recipient of the Long-Term Excellence Award from the Health Policy Statistics Section of the ASA, and elected recipient of the ISPOR Avedis Donabedian Outcomes Research Lifetime Achievement Award.

Session 4: Statistical Methods for Logistics of Futility Monitoring

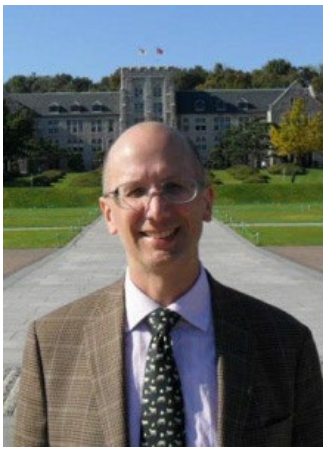
Dr. Dong-Yun Kim

National Heart, Lung, and Blood Institute

Title: Continuous Monitoring in Clinical Trials: A Fully Sequential Perspective

Abstract: In this talk, we introduce a new continuous monitoring method for the event rate of time-to-event data when patients enter the clinical trial in a random, staggered fashion. The method utilizes a fully sequential test together with newly obtained formulas for boundaries. During the trial, the continuous monitoring can detect if the empirical rate is substantially different from the target event rate. This feature is particularly useful for the cases where the target event rate is unlikely to be reached by the end of study. The accrued data can be used to answer questions that may be useful for adaptive purposes. We illustrate the method using data from a large Phase III clinical trial.

Biography: Dr. Dong-Yun Kim is a mathematical statistician at the Office of Biostatistics Research (OBR) within National Heart, Lung, and Blood Institute, National Institutes of Health in Bethesda, Maryland. She received a Ph.D. in Statistics from the University of Michigan, Ann Arbor in 2003. Before joining NIH in 2013, she held a faculty position at Virginia Tech. Her research interests include fully sequential monitoring in clinical trials, change-point inference and analysis of medical image data. Currently, she is involved in large NHLBI-sponsored clinical trials and intramural projects in MRI imaging and pulmonary diseases. Dr. Kim has years of experience in collaborative research in other areas including mobile health, bioengineering, and environmental science. She is serving as President of the Caucus for Women in Statistics and Data Science (CWS), and a board member for Korean International Statistics Society (KISS).



Dr. Daniel F. Heitjan

Southern Methodist University

Title: Real-Time Prediction in Clinical Trials: A Statistical History of REMATCH

Abstract: Randomized clinical trial designs often incorporate one or more planned interim analyses. In event-based trials, one may prefer to schedule the interim analyses at the times of occurrence of specified landmark events, such as the 100th event, the 200th event, and so on. Because an interim analysis can impose a considerable logistical burden, and the timing of the triggering event in this kind of study is itself a random variable, it is natural to seek to predict the times of future landmark events as accurately as possible.

Early approaches to prediction used data only from previous trials, which can mislead when, as commonly occurs, enrollment and event rates differ unpredictably across studies. With contemporary clinical trial management systems, one can populate trial databases essentially instantaneously. This makes it possible to create predictions from the trial data itself — predictions that are likely to be reliable and well-calibrated statistically.

This talk will describe work that some colleagues and I have done in this area. I will set the methodologic development in the context of the study that motivated our research: REMATCH, an RCT of a heart assist device that ran from 1998 to 2001 and is considered a landmark of rigor in the device industry.

Biography: Daniel F. Heitjan is Professor and Chair in the Department of Statistics & Data Science at Southern Methodist University and Professor in the Peter O'Donnell Jr. School of Public Health at the University of Texas Southwestern Medical Center. A native of Detroit, he earned a B.Sc. in Mathematics (1981), an MSc in Statistics (1984), and a Ph.D. in Statistics (1985) from the University of Chicago. He previously served on the faculties of UCLA (1985–1988), Penn State (1988–1995), Columbia (1995–2002), and the University of Pennsylvania (2002–2014). Dr. Heitjan has over 200 publications in the literature of medicine and statistics, and is an elected Fellow of the American Statistical Association (1997), the Institute of Mathematical Statistics (2012), and the Society for Clinical Trials (2017). He has served as Program Chair of the Joint Statistical Meetings (2005), Chair of the American Statistical Association's Biometrics Section (2009), and President of the Eastern North American Region of the International Biometric Society (2013). Since 2020, he has been a deputy editor of Clinical Trials. His research interests include causal modeling, the theory of inference with incomplete data, and methods in clinical biostatistics.

Dr. Yajun Mei

Georgia Institute of Technology

Title: The Kiefer-Weiss-Lorden Minimax Framework for Interim Evaluation of Efficacy and Futility

Abstract: Futility analysis offers an invaluable method to increase efficiency and decrease costs in the group-sequential design. The existing main statistical tool is the conditional power method, which is intuitively appealing but might lose statistical optimality. Here we provide a new Kiefer-Weiss-Lorden minimax framework for evaluating efficacy (rejecting the null hypothesis) or futility (accepting the null hypothesis). The main statistical idea is to extend the standard binary hypothesis testing formulation by minimizing the maximum expected sample sizes over a range of other potential hypotheses subject to the classical Type I and II error probabilities. Here we present an interesting sequential test, called 2-sequential probability ratio test (2-SPRT), that enjoys asymptotically optimal properties, and thus potentially providing an optimal test for futility analysis. If time slows, its extension to 2-CUSUM will also be discussed when there are delayed treatment effects. Hopefully this new framework and tests provide more tools and insights in real-world group-sequential designs.

Biography: Dr. Mei is a co-director of Biostatistics, Epidemiology, Research Design (BERD) at Georgia Clinical & Translational Science Alliance and the coordinator of the M.S. STAT program in the Milton Stewart School of Industrial and Systems Engineering at Georgia Tech.

Dr. Yanhong Wu

California State University

Title: Sequential Detection, Isolation, and Estimation of Common Changes in Multi-Panel Data

Abstract: In a clinical trial, timely detection of changes in patient recruitment and/or key study parameters, e.g. proportion of events, is of great practical interest to stakeholders. We first give the results on the CUSUM and Shiryayev-Roberts (S-R) procedures for sequential detection and estimation of a change in a parameter. The procedure uses a single sequence of observations that follow an exponential family which includes both discrete (Bernoulli, binomial) and continuous (exponential and normal) random variables. The adaptive CUSUM and S-R procedures are applicable when the post-change parameter is unknown. When the data are monitored in multiple panels, we assume that a common change occurs in a proportion of the panels. We propose an online sequential procedure that detects the common change, isolates the changed panels, and estimates the change time. Real data from diverse sources including a Phase III clinical trial are used for illustration.

Biography: Yanhong Wu is currently working at California State University. He graduated from the University of Toronto and did post-doctoral training at Stanford University. Previously, he has worked at the University of Alberta and the University of Michigan and visited Stanford University as a visiting scholar. He has worked on online quality control charts; risk theory; and sequential change detection, estimation, and inference after detection. He authored a monography on *Inference for Change Point and Poste Change Means after a CUSUM Test* and co-authored a textbook *on Statistical and Probability Models*. Recently, his focus has been on sequential detection, isolation, and estimation of common changes for multiple channel/panel data and sequential detection of transient signals.

Session 5: Statistical Methods for Futility Monitoring

Dr. Michael Proschan

National Institute of Allergy and Infectious Disease

Title: A Brief Review of Futility Monitoring

Abstract: This talk reviews different methods for monitoring a clinical trial for futility, including unconditional, conditional, and predictive power, beta spending functions, and predicted interval plots. Conditional and unconditional power answer two different questions: (1) How likely is a null result? (2) Will a null result be meaningful? Both questions are important when deciding whether to stop a trial for futility. In some cases, we may stop a trial that would have reached statistical significance if carried to fruition. How likely is this scenario? What is the impact of futility monitoring on power? These topics are discussed together with practical implications of futility monitoring.

Biography: Michael Proschan received his Ph.D. in Statistics from Florida State University in 1989. He has been a Mathematical Statistician in the Biostatistics Research Branch at the National Institute of Allergy and Infectious Diseases since January 2006. Prior to coming to NIAID, he spent 16 years at the National Heart, Lung, and Blood Institute. He has co-authored three books: "Statistical Monitoring of Clinical Trials: A Unified Approach," with Gordon Lan and Janet Wittes (Springer, 2006); "Essentials of Probability Theory for Statisticians," with Pamela Shaw (CRC Press, 2016); and "Statistical Thinking in Clinical Trials" (CRC Press, 2022), and is a Fellow of the American Statistical Association. Dr. Proschan is currently an Adjunct Faculty Member for the Advanced Academic Programs at Johns Hopkins University and is a former Adjunct Faculty Member in Statistics and Biostatistics at George Washington University.

Dr. Chris Jennison

University of Bath

Title: Group Sequential Designs with Early Stopping for Efficacy and Futility

Abstract: We consider the objectives of early stopping in a confirmatory, Phase 3 trial and present group sequential designs that control type I error and power. Given a maximum sample size and number of analyses, we explain how to derive designs that optimize specific objectives involving expected sample size or expected time to a decision under the null and alternative hypotheses. Comparisons with such designs show that two standard families of error spending designs are highly efficient for a range of objectives.

Our results support the use of error spending designs to create a binding or non-binding futility stopping rule. We contrast these designs with stopping rules or guidelines based on conditional power and predictive power, and note potential pitfalls for such methods.

Biography: Christopher Jennison is Professor of Statistics at the University of Bath, U.K. His Ph.D. research at Cornell University concerned the sequential analysis of clinical trials, and he has continued to work in this area for over 40 years. His book with Professor Bruce Turnbull, "Group Sequential Methods with Applications to Clinical Trials", is a standard text on this topic and is widely used by practicing statisticians. The second edition of this book, extended to cover adaptive designs, is finally nearing completion.

Professor Jennison's research is informed by experience of clinical trial analysis at the Dana Farber Cancer Institute, Boston, by service on data monitoring committees, and by a broad range of consultancy with pharmaceutical companies.

Dr. Scott Berry

Berry Consultants, LLC.

Title: The Importance of Prospective Futility Rules and the Role of Bayesian Predictive Probabilities

Abstract: A fixed sample size trial design that powers for a single effect size at 80% is rarely the right size. Many of these trials learn the answer well before the fixed sample size and many leave the clinical question unanswered. The positive aspects of right-sizing trials for success are well-studied, but as important is that a trial can also right-size itself when the trial is unlikely to answer the important clinical question. Having prospectively defined, well-vetted futility rules is critical to quality trial designs. Futility is challenging as it involves addressing the likelihood that if the trial continues, it will contribute important clinical information – the continuation is worth the resources and patient contributions to science. The Bayesian approach is ideally suited for addressing the uncertainty in the future data – as well as the uncertainty at the time of the futility analysis. The Bayesian predictive probability as a tool for futility analyses, as well as multiple examples, will be discussed.

Biography: Scott Berry is President and a Senior Statistical Scientist at Berry Consultants, LLC. He earned his Ph.D. in statistics from Carnegie Mellon University and was an Assistant Professor at Texas A&M University before co-founding Berry Consultants in 2000. He is an adjunct faculty member in the Department of Biostatistics at the University of Kansas Medical Center.

Dr. Karen Higgins

U.S. Food and Drug Administration

Title: Futility Monitoring: An FDA Statistician's Point of View

Abstract: There are different reasons to include futility monitoring in clinical trials. We often think of futility monitoring as being planned for business decisions, to limit time and money in a trial that will likely fail its primary objective. But there are also public health and ethical reasons for conducting, or not conducting, futility monitoring in clinical trials. Important considerations in futility monitoring include maintenance of trial integrity, the methods used and their operating characteristics, and the interplay with safety monitoring. This presentation will discuss some of the whys, whens, and hows of futility monitoring from the perspective of an FDA statistician.

Biography: Karen Higgins is a Supervisory Mathematical Statistician in the Office of Biostatistics in the Office of Translational Sciences, Center for Drug Evaluation and Research, FDA. She currently provides leadership to statisticians involved in the development and application of methodology used in the regulation of Rheumatology, Transplant Medicine, and Liver drug products; however, she has vast experience with clinical trials across a broad range of therapeutic areas, including infectious diseases. Throughout her career, she has tackled complex problems and is a recognized expert in non-inferiority trial designs. Most recently, she has been a thought-leader and active member of the FDA teams tasked with the evaluation of therapeutic products for the treatment of COVID-19. Dr. Higgins received her doctorate in biostatistics from the Harvard School of Public Health.

Session 6: Case Studies for Futility Monitoring

Dr. Chris Lindsell

Duke University

Title: The Role of Futility Assessments in Informative Research

Abstract: Futility is often assessed in a clinical trial to determine whether accrual should begin or continue. Methods range from formally testing a harm hypothesis to estimating the predicted probability of success. While futility analysis is not uncommon and futility is sometimes concluded, terminating accrual is controversial, and investigators often wish to continue under the presumption that there will be signals in subgroups or residual information of importance in the trial. In this talk, a brief overview of approaches to evaluating futility will be discussed using several case studies. Emphasis will be placed on realistic futility assessments that consider time, resources, and the informativeness of data.

Biography: Chris Lindsell is Director of Data Science and Biostatistics at the Duke Clinical Research Institute, and co-Chief of the Division of Biostatistics at Duke University. His focus is on the application of rigorous methods to clinical research at the intersection of acute care and public health. He has led numerous data coordinating centers supporting epidemiological studies and clinical trials, most recently the IVY Network and ACTIV-6. His expertise spans large scale data collection, predictive modeling, traditional trials, mechanistic trials, pragmatic trials, and decentralized trials. He is also a leader in learning health systems and in developing precision approaches to treatment in sepsis and septic shock.

Dr. Vlad Dragalin

Janssen Pharmaceuticals

Title: The Utility of Futility Rules: An Industry Perspective

Abstract: For industry, the main utility of futility rules is to allow early stopping of a study when it seems unlikely to achieve its primary efficacy objectives, and it is mainly motivated by financial and ethical considerations: substantial time reduction and resource savings that could be spent on more promising research, and preventing patients from being exposed to ineffective experimental treatments unnecessarily.

After an introduction of the key concepts and brief overview of the classical statistical approaches that have been used to assess futility, including rules based upon conditional power, predictive probability, and beta spending functions, we will focus on newly proposed methods. These include futility rules based on short- and long-term endpoints, optimality considerations for competing futility rules, enhancing the futility rule by adding external clinical observations and shifting market conditions, despite meeting the clinical endpoints.

We close with a discussion of several issues to be considered in designing and executing a clinical trial with a futility stopping rule and provide some advice and perspectives that may be helpful when futility rules are considered and implemented especially compared to futility boundaries commonly applied in practice.

Biography: Vlad Dragalin is a Vice President, Scientific Fellow in Statistics and Decision Sciences at Janssen Pharmaceutical Companies of Johnson & Johnson. Vlad is the Global Head of QS Consulting and is chairing the QSTEPS (QS Technical Excellence and Program Strategy) Advisory Committee.

He joined Janssen in 2014 as the Head of the Adaptive Clinical Trial Center of Excellence. Vlad is a well-known adaptive design expert with 35 years of experience in developing the statistical methodology of adaptive designs and with more than 20 years of experience in pharmaceutical industry, including positions of increasing responsibilities at GlaxoSmithKline, Wyeth, Pfizer, Quintiles, and Aptiv Solutions.

Prior to joining the pharmaceutical industry in 1999, Vlad had a record of distinguished service for more than 15 years in various positions at prestigious research institutions in the United States (University of Rochester, NY), Moldova (institute of Mathematics, Academy of Sciences), Italy (Institute for Applied Mathematics and Informatics, Milan), and Germany (University of Wurzburg). Vlad received his Ph.D. in Probability Theory and Mathematical Statistics from the Steklov Mathematical Institute, USSR Academy of Sciences, Moscow in 1988.

Vlad is a Fellow of the American Statistical Association and an Associate Editor of the Journal of Biopharmaceutical Statistics and Statistics in Biopharmaceutical Research. Vlad is a PhRMA representative on ICH E9(R1) Expert Working Group on Estimands and Sensitivity Analyses and on ICH E20 Expert Working Group on Adaptive Designs. He is a member of PhRMA Clinical Development Workgroup.

Dr. Erica H. Brittain

National Institute for Allergy and Infectious Diseases

Title: Perspectives on Futility Analysis in Practice

Abstract: This talk will focus on practical issues in futility analyses and touch on several real-world examples. Topics discussed will include a) unplanned futility analyses, b) small trials, and c) consideration of more than the primary endpoint.

Biography: Erica Brittain is Deputy Branch Chief for Biostatistics Research at the National Institute for Allergy and Infectious Diseases (NIAID). Since arriving at NIAID in 2003, she has focused on trial design issues in allergy and autoimmune diseases, and in recent years, COVID. In her current role, she has advised on hundreds of trials. Prior to NIAID, she worked at FDA, Statistics Collaborative, and NHLBI. In addition, she has served on numerous advisory panels for FDA and NIH. She has published on a wide array of issues in clinical trial design, including some of the earliest work in adaptive designs. Other topics include non-inferiority designs, sequential monitoring, factorial designs, survival analysis, and causal inference. She recently co-authored a book on statistical hypothesis testing in the context of science. She has also served as an Associate Editor of *Controlled Clinical Trials* (the predecessor of *Clinical Trials*). She received her M.S. in Statistics from Stanford in 1980, and her Ph.D. in Biostatistics from the University of North Carolina in 1984.

Dr. Wesley Self

Vanderbilt University Medical Center

Title: ACTIV-4 Host Tissue Trial: Design and Interim Monitoring

Abstract: In this presentation, we will discuss the design of the ACTIV-4 Host Tissue platform (NCT04924660), with an emphasis on interim monitoring. This NHLBI-funded platform trial is evaluating the efficacy and safety of multiple agents for severe COVID-19 among adults hospitalized with COVID-19 associated hypoxemia. The first two trials on the platform were halted based on interim monitoring. The discussion will focus on why the interim monitoring parameters were selected and how they performed in the trials.

Biography: Wesley H. Self, M.D., M.P.H., is the Senior Vice President for Clinical Research and Principal Investigator for Clinical and Translational Science Award at Vanderbilt University Medical Center. He is a physician who treats patients in the emergency department setting and clinical researcher who focuses on late-stage clinical trials and effectiveness studies. He is a leader of several multicenter clinical research networks, including the NHLBI PETAL Network and ACTIV-4 Host Tissue Platform, the NIAID STRIVE Network, the CDC IVY Network, and DoD Pragmatic Critical Care Research Group. His work has consistently resulted in important findings for patient care, including in the fields of sepsis resuscitation (NEJM 2023), respiratory failure (NEJM 2022), COVID-19 vaccines (JAMA 2021), appendicitis (NEJM 2020), COVID-19 monoclonal antibody therapeutics (NEJM 2021), COVID-19 host response therapeutics (JAMA 2020), ARDS (NEJM 2019), intubation techniques (NEJM 2019), intravenous fluid management (NEJM 2018), influenza vaccines (JAMA 2015), and pneumonia epidemiology (NEJM 2015). Dr. Self graduated from medical school at the University of Virginia, received his clinical training at Northwestern University, and his research training at Vanderbilt University. He currently leads a clinical and translational research portfolio at Vanderbilt.

Dr. Nancy L. Geller

National Heart, Lung, and Blood Institute

Biography: Nancy L. Geller is the Director of the Office of Biostatistics Research of the National Heart, Lung and Blood Institute (NHLBI). She directs a group of 12 statisticians who collaborate in the design, implementation, monitoring and analysis of clinical studies in heart, lung and blood diseases and sleep disorders and administers all statistical activities of the NHLBI. She is involved in the design and analysis of a number of cardiovascular and hematology trials, including the Chronic Hypertension and Pregnancy trial (CHAP), a Family-Based Behavioral Treatment versus Usual Care for Childhood Obesity study, known as PLAN (Primary care Pediatrics, Learning And Nutrition randomized trial), and a number of Bone Marrow Transplantation Clinical Trial Network trials. She has published over 300 papers in the biostatistical and medical literature. She is an Associate Editor of Biometrics and Fellow of the International Statistics Institute, the American Statistical Association and the Society for Clinical Trials. She was President of the American Statistical Association in 2011.