Short Courses of ISCA 2020 Applied Statistics Symposium

Multivariate meta-analysis methods

Haitao Chu, MD, PhD, Professor of Biostatistics, University of Minnesota Twin Cities

Yong Chen, PhD, Associate Professor of Biostatistics, University of Pennsylvania

Duration: Half Day - morning

Abstract:

Comparative effectiveness research aims to inform health care decisions concerning the benefits and risks of different prevention strategies, diagnostic instruments and treatment options. A meta-analysis is a statistical method that combines results of multiple independent studies to improve statistical power and to reduce certain biases compared to individual studies. Meta-analysis also has the capacity to contrast results from different studies and identify patterns and sources of disagreement among those results. The increasing number of prevention strategies, assessment instruments and treatment options for a given disease condition have generated a need to simultaneously compare multiple options in clinical practice using rigorous multivariate meta-analysis methods.

This short course, co-taught by Drs. Chu and Chen who have collaborated on this topic for more than a decade, will focus on most recent developments for multivariate meta-analysis methods. This short course will offer a comprehensive overview of new approaches, modeling, and applications on multivariate meta-analysis. Specifically, the instructors will discuss the contrast-based and arm-based network meta-analysis methods for multiple treatment comparisons; network meta-analysis methods for multiple diagnostic tests; and multivariate meta-analysis methods estimating complier average causal effect in randomized clinical trials with noncompliance.

Case studies will be used to illustrate the principles and statistical methods introduced in this course. This application oriented short course should be of interest to researchers who would apply up-to-date multivariate meta-analysis methods. We anticipate that it will be well-received by an interdisciplinary scientific community, and play an important role in improving the rigor and broadening the applications of multivariate meta-analysis.

About the instructors:

Dr. Chu is Professor of Biostatistics at University of Minnesota Twin Cities. He is an ASA Fellow and elected member of the Society for Research Synthesis Methodology since 2016. Dr. Chu's research lies at the intersection of biostatistics and epidemiology, with a recent focus on multivariate research synthesis methods. Dr. Chu has published over 170 peer-reviewed articles

with over 10,000 Google Scholar citations. Specifically, Dr. Chu has published over 50 peer-reviewed manuscripts on systematic reviews and meta-analysis in top ranked statistical and medical journals such as JASA, Biometrics, Biostatistics, SIM, SMMR, BMJ, Clinical Trials, JNCI, AIDS, Epidemiology and AJE. Dr. Chu's research on innovative statistical methods improve meta-analysis has been supported by 8 grants from FDA, AHRQ, NIAID, NIDCR and NLM as the principal investigator. Dr. Chu serves as an Associate Editor for Journal of the American Statistical Association, the American Journal of Epidemiology, and Statistics and Its Interface.

Dr. Yong Chen is Associate Professor of Biostatistics at University of Pennsylvania. He directs a Computing, Inference and Learning Lab at University of Pennsylvania, which focuses on integrating fundamental principles and wisdoms of statistics into quantitative methods for tackling key challenges in modern biomedical data. Dr. Chen is an expert in synthesis of evidence from multiple data sources, including systematic review and meta-analysis, distributed algorithms, and data integration, with applications to comparative effectiveness studies, health policy, and precision medicine. He is also working on developing methods to deal with suboptimal data quality issues in health system data, dynamic risk prediction, pharmacovigilance, and personalized health management. He has over 100 publications in a wide spectrum of methodological and clinical areas. He has been principal investigator on a number of grants, including R01s from the National Library of Medicine and National Institute of Allergy and Infectious Diseases, and Improving Methods for Conducting Patient-Centered Outcomes Research grant from Patient-Centered Outcomes Research Institute. He is an elected fellow of the Society for Research Synthesis Methodology, and the International Statistical Institute.

Including historical data in clinical trial design and analysis

Frank Fleischer, PhD, Head of Methodology Statistics, Boehringer-Ingelheim Pharma GmbH & Co. KG

Martin Oliver Sailer, PhD, Methodology Expert Statistician, Boehringer-Ingelheim Pharma GmbH & Co. KG

Duration: Half Day – morning

Abstract:

With the growing number of targeted drug development programs, there is an ever increasing interest to make these programs more cost effective. Borrowing of information from historical data allows to reduce the number of patients recruited to new trials and helps to bring new therapies to patients faster. Participants will learn requirements for the use of historical data in clinical trial design and analysis. It will be shown how Bayesian hierarchical models can be used to borrow information from historical data and perform Bayesian evidence synthesis with meta-analytic predictive priors. Advantages of dynamic weighting will be motivated. Since the population in the historical data and the new study may differ, propensity score methods and methods for covariate adjustment need to be considered. Case studies will be presented for examples from dose finding in oncology, basket trials and go/no-go decision making after phase II. Considerations for confirmatory settings will be addressed. Participants will be able to implement methods with computer exercises.

About the instructors:

Dr. Frank Fleischer

Being a trained mathematician and statistician Frank has worked for more than 10 years in the pharmaceutical industry. He is heading a global team of statisticians at Boehringer Ingelheim focusing on statistical methodology and the implementation of innovative statistical designs into practice. In that role, Frank and his team are considered with methodological questions regarding adaptive designs, statistical decision making, dose finding and Bayesian borrowing designs as well as with piloting these methods in clinical trials. Through this function several projects across different therapeutic areas and phases are supported. Formerly he has been a lead project statistician for different projects in oncology, immunology and the biosimilars.

Dr. Martin Oliver Sailer,

With nine years of experience in the pharmaceutical industry, he has been Statistical lead for multiple pivotal Oncology and Biosimilar development programs. His consulting work focuses on introducing Bayesian methods in all phases of clinical development. His research interests include Design of Experiments, Bayesian Statistics, Basket designs, Statistical Go/No-Go decision making, and Estimands. He studied Statistics at TU Dortmund University in Germany and Iowa State University, Ames, IA.

Short Course on Absolute Risk Prediction

Mitchell H Gail, MD, PhD, Senior Investigator Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH trained at Harvard Medical School and in statistics at George Washington University. He is an Adjunct Professor in the Department of Biostatistics, Johns Hopkins University

Ruth Pfeiffer, PhD, Senior Investigator, Biostatistics Branch, National Cancer Institute, graduate of the University of Maryland, College Park and the Technical University of Vienna, Austria
Duration: Full Day

Abstract:

Absolute (or "crude") risk is the probability that an individual who is free of a given disease at an initial age, a, will develop that disease in the subsequent interval (a, t]. Absolute risk is reduced by mortality from competing risks. Models of absolute risk that depend on covariates have been used to design intervention studies, to counsel patients regarding their risks of disease and to inform clinical decisions. This course will define absolute risk and discuss methodological issues relevant to the development and evaluation of risk prediction models. Various study designs and data for model building will be presented, including cohort, nested case-control, and case-control data combined with registry data. Issues relating to the evaluation of risk prediction models and the strengths and limitations of risk prediction models for various applications will be discussed. Standard criteria for model assessment will be presented, as well as loss function-based criteria applied to the use of risk models to screen a population and the use of risk models to decide whether to take a preventive intervention that has both beneficial and adverse effects. Methods for validating models in independent data when some predictors are missing are presented. Finally, updating risk models when information on new predictors becomes available will be discussed.

About the instructors:

Dr. Mitchell H. Gail is a Senior Investigator at the Biostatistics Branch of the Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI). Dr. Gail's current research interests include statistical methods for the design and analysis of epidemiologic studies, and the development and application of models to predict the absolute risk of disease. Dr. Pfeiffer and Dr. Gail recently wrote a book entitled "Absolute Risk: Methods and Applications in Clinical Management and Public Health". Dr. Gail served as President of the American Statistical Association and is a member of the National Academy of Medicine.

Dr. Ruth Pfeiffer is a tenured senior investigator at the Biostatistics Branch of the Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI). She received an M.S. degree in applied mathematics from the Technical University of Vienna, Austria, an M.A. degree in applied statistics and a Ph.D. in mathematical statistics both from the University of Maryland, College Park. At NCI she is an active collaborator on many research projects and mentors several fellows and junior investigators. Her research focuses on statistical methods for absolute risk prediction, problems arising in molecular and genetic epidemiologic studies and the analysis of data from electronic medical records. She is the recipient of a Fulbright

Fellowship, an elected Member of the International Statistical Institute, and an elected Fellow of the American Statistical Association.

Utilizing Real-World-Data and Real-World-Evidence in Drug Development and Evaluation

Binbing Yu, PhD, Associate Director, AstraZeneca Oncology Biometrics

Bo Lu, PhD, Professor, Ohio State University (OSU), Division of Biostatistics College of Public Health

Qing Li, PhD, Senior Statistician, Takeda Pharmaceutical Company

Duration: Full Day

Abstract:

In recent years, the rapid increase in the volume, variety, and accessibility of digitized RWD and RWE has presented unprecedented opportunities for the use of RWD and RWE throughout the drug product lifecycle. In clinical development, RWD and RWE have the potential to improve the planning and execution of clinical trials, and create a virtual control arm for a single arm for accelerated approval and label expansion. From the product lifecycle perspective, effective insights gleaned from RWE bring about informative relative benefits of drugs, comparative effectiveness, price optimization, and new indications.

The goal of the short course is to serve as resources for practitioners who wish to apply these modern statistics and analytics in drug research and development. This short course will cover the essential statistical methodology for causal inference and recent practical case studies that adopted RWD and RWE in the clinical development and evaluation. In the morning session, we will introduce the current trend, challenges and opportunities of RWD and RWE in drug development and evaluation. We will also provide a comprehensive review of the relevant statistical methods for treatment effect estimation using non-randomized data, including propensity score matching/stratification/weighting and sensitivity analysis. In the afternoon session, we will illustrate how to apply advanced statistical tools to practical case studies, including RWD and RWE in the clinical development and post-marketing drug development.

About the instructors:

Dr. Binbing Yu is an Associate Director in the Oncology Statistical Innovation group in AstraZeneca. He serves as the statistical expert across the whole spectrum of drug R&D process, including early-clinical and clinical research, design, operation and manufacturing, clinical pharmacology, oncology medical affairs and post-marketing surveillance. He obtained his PhD in Statistics from the George Washington University. His primary research interests are clinical trial design and analysis, cancer epidemiology, cause inference in observation studies, PK/PD modeling and Bayesian analysis. He was previously the Biometry Section Chief in the National Institute on Aging. He has nearly 80 publications in scientific and statistical journals and published a book on statistical methods on immunogenicity.

Dr. Bo Lu is a Professor of Biostatistics in the College of Public Health, the Ohio State University. He obtained his PhD in Statistics from the University of Pennsylvania. His primary research interest covers causal inference with observational data, matching/weighting adjustment for complex designs including multiple treatment arms, time-varying treatment initiation, with complex survey weights, etc. Bayesian nonparametric modeling for heterogeneous causal effects, and statistical methods for survey sampling. He has been PIs for both federal and local-funded research grants on causal inference methodology. He has served as the lead statistician for the Ohio Medicaid Assessment Survey series since 2008. He also has extensive collaborations with Pharmaceutical industry on utilizing causal inference methods to leverage RWD in drug discovery.

Dr. Qing Li is a senior statistician in the statistical methodology group under the statistics and quantitative science (SQS) department at Takeda Pharmaceutical Company. His responsibilities include statistical methodology development and consultation for real-world-evidence (RWE) and advanced adaptive design from proof-of-concept to late phase studies across multiple therapeutic areas including oncology, gastroenterology (GI), rare disease, and vaccine. His research interests include propensity score (PS) methods, RWE, adaptive designs (sample size re-estimation, subgroup enrichment design, seamless design), Immuno-Oncology (IO) design and surrogate endpoints. He obtained his MS and PhD degree in biostatistics from the University of Iowa.

Empower Statistician with Spark, Machine Learning and Deep Learning

Hui Lin, PhD, Head of Data Science at Netlify

Ming Li, PhD, Research Scientist at Amazon

Duration: Full Day

Abstract:

Data can be a valuable asset, especially when there's a lot of it. Exploratory data analysis, business intelligence, and machine learning can benefit tremendously if such big data can be wrangled and modelled at scale. Apache Spark is an open-source distributed engine for querying, processing and modeling big data. In this one-day workshop, you will learn how to leverage Spark and R/Python to process and model big data with common machine learning algorithm. By the end of this workshop, you will have a solid understanding of how to process big data using Spark and how to build common machine learning models in the cloud environment. You will also learn the motivation and use cases of deep learning through hands-on exercises. This workshop is designed for audience with statistics education background. This course bridges the gap between traditional statisticians and data scientists. No software download or installation is needed, everything is done through laptop's internet browser (Chrome or Firefox) with Databricks free cloud environment.

About the instructors:

Hui Lin is the head of data science at Netlify where she is leading and building the data science department. Before Netlify, she was a Data Scientist at DuPont. She provided data science leadership for a broad range of predictive analytics and market research analysis from 2013 to 2018. She is the co-founder of Central Iowa R User Group, blogger of https://scientistcafe.com/, and 2018 Program Chair of ASA Statistics in Marketing Section. She enjoys making analytics accessible to a broad audience and teaches tutorials and workshops for practitioners on data science (https://course2019.scientistcafe.com/). She holds MS and Ph.D. in statistics from Iowa State University.

Dr. Ming Li is currently a Research Scientist at Amazon. He organized and presented 2018 JSM Introductory Overview Lecture: Leading Data Science: Talent, Strategy, and Impact. He was the Chair of Quality & Productivity Section of ASA. He was a Data Scientist at Walmart and a Statistical Leader at General Electric Global Research Center before joining Amazon. He obtained his Ph.D. in Statistics from Iowa State University in 2010. With deep statistics background and a few years' experience in data science and machine learning, he has trained and mentored numerous junior data scientist with different backgrounds such as statistician, programmer, software developer, database administrator and business analyst. He is also an Instructor of Amazon's internal Machine Learning University and was one of the key founding members of Walmart's Analytics Rotational Program.

Estimands and Statistical Methods for Missing data in Clinical Trials

Frank Liu, PhD, Distinguished Scientist, Merck & Co.

Mandy Jin, PhD, Director of Clinical Statistics, AbbVie Inc.

Duration: Half Day - afternoon

Abstract:

In longitudinal clinical trials, data may be missing due to intercurrent events such as missing visits or early discontinuation. The strategies discussed in ICH E9 (R1) addendum for handling intercurrent events requires clearly defined estimands and associated assumptions about missing data. To evaluate the underline treatment effects of an investigational new drug or biologics, it is desirable to consider estimands that can define an attributable causal inference for outcomes. Properly analyzing missing data with appropriate methods is critical to assess the attributable estimands.

Commonly used approaches for missing data assume data are missing at random (MAR) and analyze data using likelihood-based methods or multiple imputations (MI). Because the MAR assumption is often difficult to justify, both regulatory agencies and industry sponsors have been seeking alternative approaches to handle missing data under missing not at random (MNAR) assumption, which estimates attributable estimands while excluding potential confounding.

This half-day tutorial is intended to cover various methods that have been advocated in dealing with missing data and illustrates how to implement the analyses methods using examples. The tutorial begins with a review of estimands associated with missing data, followed by an overview of conventional methods for missing data handling such as maximum likelihood methods, multiple imputation, generalized estimation equation approaches, and Bayesian methods. The rest of the course is devoted to recently developed methods, including control-based imputation, tipping point analysis, and some methods developed by the instructors. Real clinical trial examples will be presented for illustration with implementation of the analysis using SAS software, including the MIXED, MI, MIANALYZE, GEE, and MCMC procedures.

About the instructors:

Dr. G. Frank Liu is a distinguished scientist at Merck & Co., Inc. and a Fellow of the American Statistical Association (ASA). For more than 24 years at Merck, Frank has gained extensive industry working experiences. His research interests include methods for longitudinal trials, missing data, safety analysis, and noninferiority trials; and has published more than 40 peer-reviewed statistical papers. He has been leading the development of many methodological guidance documents within Merck. He has taught short courses previously at Deming conferences, Biopharmaceutical Regulatory-Industry workshops, ASA conference on statistical practice, and conferences of the International Society of Biopharmaceutical Statistics.

Dr. Mandy Jin is currently a Director of Clinical Statistics at AbbVie Inc. She has gained 12 years of experience in clinical research across different therapeutic areas since she obtained her PhD in statistics from Columbia University in 2008. Her research interests include statistical methodologies for clinical trials, such as missing data, Bayesian analysis, adaptive designs, multiplicity adjustment, and machine learning. She has published more than 20 peer-reviewed statistical papers in these topics.

Statistical Remedies for Flawed Conventions in Medical Research

Peter F. Thall, PhD, Department of Biostatistics, The University of Texas MD Anderson Cancer Center

Duration: Half Day - afternoon

Abstract:

Many statistical methods commonly used for data analysis or clinical trial design by medical researchers are severely flawed. Unfortunately, some of these dysfunctional statistical conventions and paradigms are deeply embedded in the medical research community, and have become standard or even required practice. Ultimately, the consequence is that practicing physicians are misled to choose inferior or even harmful treatments for their patients. In this half day short course, I will identify and describe, by example, severe problems with a variety of statistical practices commonly used by medical statisticians and physician researchers. For each flawed practice, I will provide at least one practical alternative. Topics to be covered will include misinterpreting tests of hypotheses, misuse of p-values, evaluating strength of evidence, relationships between early treatment response and survival time, being misled by single-arm trials, futile futility rules, unsafe safety rules, Simpson's paradox, biomarkers and stratification, randomization and causality, bias correction, problems with outcome adaptive randomization, cherry picking, phase II-III designs, and dynamic treatment regimes.

About the instructor:

Dr. Peter F. Thall is the Anise J. Sorrell Professor in the Department of Biostatistics at M.D. Anderson Cancer Center. He is a Fellow of the American Statistical Association (ASA) and the Society for Clinical Trials, received the Don Owen Award in 2014, and is an ASA media expert. Dr. Thall has published over 260 papers and book chapters in the statistical and medical literature, and co-authored the 2016 book Bayesian Designs for Phase I-II Clinical Trials. His latest book, Statistical Remedies for Medical Researchers will published in early 2020. Dr. Thall's research areas include clinical trial design, precision medicine, Bayesian nonparametric statistics, incorporating expert opinion into Bayesian inference, and dynamic treatment regimes. He has presented over 200 invited talks and 30 short courses, and served as an associate editor for Journal of the National Cancer Institute, Statistics in Medicine, Statistics in Biosciences, Clinical Trials, and Biometrics.

Statistics and Machine Learning Methods for EHR Data: From Data Extraction to Data Analytics/Predictions

Hulin Wu, PhD, The Betty Wheless Trotter Professor and Chair, Department of Biostatistics & Data Science, School of Public Health, University of Texas Health Science Center at Houston (UTHealth), Director, Center for Big Data in Health Sciences

Vahed Maroufy, PhD, Assistant Professor, Department of Biostatistics & Data Science, School of Public Health, University of Texas Health Science Center-Houston (UTHealth)

Ashraf Yaseen, PhD, Assistant Professor, Department of Biostatistics & Data Science, School of Public Health, University of Texas Health Science Center-Houston (UTHealth)

Duration: Full Day

Abstract:

This short course will provide an overview and present details of electronic health record (EHR) data extraction, cleaning, processing and analytics for scientific discoveries. The use of EHR data is becoming more prevalent for research purpose and deriving real-world evidence for decision or policy-making. However, analysis of this type of data has many unique complications due to how they are collected, processed, missing data issues, and types of questions that can be answered. This proposed short course covers many important topics related to using EHR data for research and scientific discoveries that include data extraction, cleaning, processing, making inference, and predictions based on many years of practical experience of instructors and their collaborators in the EHR Working Group at the University of Texas Health Science Center at Houston (UTHealth). Statistical and machine learning approaches will also be presented for EHR data extraction, cleaning and analysis. Additionally, since research projects for EHR Big Data are being conducted in large multidisciplinary research groups, the approaches for multiple-project management are necessary and will be also covered in this course.

About the instructors:

Dr. Wu joined the University of Texas Health Science Center at Houston (UTHealth) as Dr. D.R. Seth Family Professor and Associate Chair of Biostatistics and Professor of Biomedical Informatics in September 2015. He was appointed as the endowed Betty Wheless Trotter Professor and Chair for the newly named Department of Biostatistics & Data Science, UTHealth School of Public Health (SPH) in 2017. He is the Founding Director of the "Center for Big Data in Health Sciences" at UTHealth SPH with a goal to develop and use cutting-edge data science approaches to deal with Big Data from biomedical and health sciences. Dr. Wu was Dean's Professor of Biostatistics and Computational Biology, Professor of Medicine, and Professor of Public Health Sciences at the University of Rochester Medical Center (URMC) from 2003-2015. He was the URMC Founding Director of the Center for Integrative Bioinformatics and Experimental Mathematics. Dr. Wu has extensive experience in directing NIH-funded research

projects and contracts. As PI/Co-PI, he has been continuously funded by NIH since 1998 and he has received a total of $30 million in NIH funding for independent research (R29 and 5 R01 grants), T32 training grant and NIH Cooperative Contract or center grants in the past 20 years. Dr. Wu has published 2 books and more than 130 peer-reviewed papers in statistics/biostatistics, biomathematics, bioinformatics and biomedical journals.

Dr. Maroufy is an Assistant Professor of Biostatistics at the department of Biostatistics and Data Science, School of Public Health-UTHealth. His research interests include data mining, statistical analysis and predictive modeling using big Electronic Health Records (EHR) and claim datasets. Currently his focus is on EHR data processing, cleaning, missing imputation and predictive analysis. Dr. Maroufy, has also experience and expertise in mathematical and methodological statistics such as mixture models, measurement error and sensitivity analysis using high-dimensional data.

Dr. Yaseen is currently an Assistant Professor of Data Science at the School of Public Health-UTHealth. His research interests include Machine Learning, Data Management & Analysis, Big Data, Bioinformatics, and High Performance Computing. In his current research work, Dr. Yaseen is exploring Big Data and Deep Learning technologies in Electronic Health Records data to address clinical and public health questions. He has extensive experience in computer programming, database design, implementation and management, web design and programming, and software engineering. He is actively contributing to several research projects at UTHealth for health-data analysis.

Statistical Analysis of Microbiome Data with R

Yinglin Xia, PhD, Research Associate Professor, University of Illinois at Chicago

(Din) Ding-Geng Chen, PhD, Wallace H. Kuralt Distinguished Professor, University of North Carolina at Chapel Hill

Duration: Full Day

Abstract:

Microbiome data are generated through either 16S rRNA gene sequencing or shotgun metagenomic sequencing. One unique feature of microbiome data is phylogenetic tree-structured. The bacterial taxa in a community are not randomly distributed; they usually not only depend on each other, but also exist the phylogenetic relationships among bacteria, which provides insights into the evolutionary relationships among bacterial taxa: a phylogenetic tree. Microbiome data have several features. The taxa abundance, amplicon sequence variants (ASVs) or operational taxonomic unit (OTU) counts, are naturally constrained, high dimensional, sparse with containing a large proportion of zero counts in the analysis data: feature table or OTU table. Typically, these data have complex covariance and correlation structures among different ASVs, OTUs, or taxa, and over-dispersed with large within-group heterogeneities.

The unique data structure and all these data features pose the great challenges to analyze microbiome data using standard statistical methods and models. Recently we developed a statistical framework which consists of combining newly developed methods and models for microbiome data and borrowing methods and models from other fields such as ecology. This work was published in 2018 as a book titled “Statistical Analysis of Microbiome Data with R” by Springer (coauthored by Xia, Y., Sun, J. and Chen, D.G.) (https://www.springer.com/us/book/9789811315336). Since the book published in October, 2018, there are more than 40,000 downloads from Springer Bookmetrix, which is far more than the average downloads of statistical book from Springer. So far the readers from more than 30 countries have given us feedbacks and we were told that this book has been used as textbook in Japan and several US universities. We were contacted frequently for requesting book material and slides for their teaching. The book review editor of the Biometrical Journal (Prof. and Dr. Annette Kopp-Schneider, the head of Division of Biostatistics, German Cancer Research Center, Germany) solicited a book review of this book, which published on 21 June 2019. This book was very positively reviewed (Biometrical Journal. 2019;1–2. www.biometrical-journal.com © 2019 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim DOI: 10.1002/bimj.201900176).

Given the importance of microbiome study and currently only statistical book available, this book has been well received by peers of microbiome research. This course is designed to use this new book in this ICSA conference to meet the need of students and faculty to understand the microbiome data and perform the statistical analysis of microbiome data with R.

About the instructors:

Dr. Yinglin Xia is a Research Associate Professor at the Department of Medicine, the University of Illinois at Chicago, USA. He was a Research Assistant Professor in the Department of Biostatistics and Computational Biology at the University of Rochester, Rochester, NY. Dr. Xia has worked on a variety of research projects and clinical trials in microbiome, gastroenterology, oncology, immunology, psychiatry, sleep, neuroscience, HIV, mental health, public health, social and behavioral sciences, as well as nursing caregiver. He has published more than 100 papers in peer-reviewed journals on Statistical Methodology, Clinical Trial, Medical Statistics, Biomedical Sciences, and Social and Behavioral sciences. He serves the editorial board for several scientific journals. He has successfully applied his statistical knowledge, modeling and programming skills to study designs and data analysis in biomedical research, clinical trials, and in microbiome research. He has written the first book, an invited review, and a book chapter on statistical analysis of microbiome data. He has designed four grants on microbiome studies funded by NIH, VA, and other funding agencies. His recent papers on microbiome data analysis are well received by peers.

Dr. Din Chen is a Fellow of ASA. He is now the Wallace H. Kuralt distinguished professor in biostatistics, University of North Carolina at Chapel Hill. He was a professor in biostatistics at the University of Rochester and the Karl E. Peace endowed eminent scholar chair in biostatistics at Georgia Southern University. Professor Chen is also a senior statistics consultant for biopharmaceuticals and government agencies with extensive expertise in clinical trials and bioinformatics. He has more than 150 referred professional publications and co-authored/co-edited 23 books on randomized clinical trials, statistical meta-analysis, public health statistical methods, causal inferences and statistical Monte-Carlo simulation and public health applications.