# ENAR 2016 Short Courses

| | REGISTRATION RECEIVED | | | | | |
|---|---|---|---|---|---|---|
| | By Feb.1 | | | After Feb.1 | | |
| | Half Day | Second Half Day | Full Day | Half Day | Second Half Day | Full Day |
| **Member** | $225 | $190 | $325 | $250 | $215 | $350 |
| **Non-Member** | $275 | $240 | $375 | $300 | $265 | $400 |

Sunday March 6

## SC1: Missing Data in Regression Models

**FULL DAY: 8:00 am to 5:00 pm**

**Joe Ibrahim**
University of North Carolina at Chapel Hill

**Overview:** Missing data is a major issue in many applied problems, especially in the biomedical sciences. In this short course, we present a detailed account of the four common paradigms for inference in missing data problems. These are i) Maximum Likelihood (ML), ii) Multiple Imputation (MI), iii) Fully Bayesian (FB), and Weighted Estimating Equations (WEE). There is considerable interest as to how these four methodologies are related, the properties of each approach, the advantages and disadvantages of each methodology, and their computational implementation. We also discuss the various assumptions and definitions of missing data mechanisms, present several examples, and examine data that are missing at random (MAR) and/or nonignorably missing (MNAR), as well as missing covariate and/or response data. We will also discuss methods and applications for several types of models including generalized linear models, models for longitudinal data, and survival models. For each paradigm (ML, MI, FB, WEE) and model, we will present detailed case studies and software implementation in SAS, WinBUGS, and Cytel Studio's XMISS package.

**Instructor Biography:** Dr. Joseph Ibrahim is Alumni Distinguished Professor, Director of Graduate Studies, and Program Director of the Cancer Genomics Training grant, in the Department of Biostatistics at UNC. His areas of research focus are Bayesian inference, missing data problems, clinical trials, and cancer genomics. He has directed or co-directed 25 doctoral students and 8 post-doctoral fellows. He has taught courses in Bayesian statistics, Advanced Statistical Inference, Generalized Linear Models, and Missing Data in the Department of Biostatistics at UNC. He is currently the Editor for the Journal of the American Statistical Association – Applications and Case Studies. Dr. Ibrahim has published over 265 research papers, mostly in top statistical journals. He also has published two advanced graduate-level books on Bayesian survival analysis and Monte Carlo methods in Bayesian computation. Dr. Ibrahim has a long standing history of teaching half day and full day short courses at ENAR, JSM, and at pharmaceutical companies. He has done full day short courses in Meta-analysis and Network Meta-analysis, Bayesian Methods for Computation, Bayesian Survival Analysis, Missing Data in Regression Models, Bayesian Methods in SAS, Bayesian Methods in Clinical Trials, Joint Modeling of Longitudinal and Survival Data, Introduction to Longitudinal Data Modeling, Introduction to Bayesian Methods, and Informative Prior Elicitation. He is an elected fellow of the American Statistical Association and the Institute of Mathematical Statistics, and an elected member of the International Statistical Institute.

## SC2: Statistical Analysis of Network Data

**FULL DAY: 8:00 am to 5:00 pm**

**Eric Kolaczyk**
Boston University

**Overview:** Networks have permeated everyday life through everyday realities like the Internet, social networks, and viral marketing. Their use has become especially prevalent in the biological and life sciences, particularly in computational biology and neuroscience. Accordingly, network analysis is an important growth area in the quantitative sciences, with roots in social network analysis going back to the 1930s and graph theory going back centuries. Measurement and analysis are integral components of network research, and statistical methods therefore play a critical role in network analysis. This course will provide a broad treatment of foundational topics relevant to statistical analysis of network data across the disciplines. Material will be organized according to a statistical taxonomy, with presentation entailing a conscious balance of conceptual and technical aspects. Additionally, practical application of network analysis will be demonstrated in the context of the R software environment. Topics for the morning will include manipulation, visualization, and descriptive analysis of network data. In the afternoon, we will focus on network sampling and inference, and the modeling of networks and network-indexed processes. Specific examples of network analysis will be drawn from a variety of domain areas, with emphasis on computational biology and neuroscience and on social networks.

**Instructor Biography:** Eric Kolaczyk is Professor of Statistics, and Director of the Program in Statistics, in the Department of Mathematics and Statistics at Boston University, where he also is an affiliated faculty member in the Program in Bioinformatics, the Program in Computational Neuroscience, and the Division of Systems Engineering. Prof. Kolaczyk's main research interests currently revolve around the statistical analysis of network-indexed data, and include both the development of basic methodology and inter- disciplinary work with collaborators in bioinformatics, computer science, geography, neuroscience, and sociology. Besides various research articles on these topics, he has also authored two books in this area— Statistical Analysis of Network Data: Methods and Models (Springer, 2009) and Statistical Analysis of Network Data with R (Springer, 2014), joint with Gabor Csardi. He has given various short courses on material from his book in recent years, including for the Center for Disease Control (CDC) and the Statistical and Applied Mathematical Sciences Institute (SAMSI) in the US, as well as similar venues in Belgium, England, and France. Prof. Kolaczyk has served as associate editor on several journals, including currently the Journal of the American Statistical Association and the IEEE Transactions on Network Science and Engineering. He has also served as (co) organizer for workshops focused on networks and network data, including as lead organizer for a year-long program at SAMSI in 2010-11. He is an elected fellow of the American Statistical Association (ASA), an elected senior member of the Institute for Electrical and Electronics Engineers (IEEE), and an elected member of the International Statistical Institute (ISI).

## SC3: Introduction to Statistical Machine Learning

**FULL DAY: 8:00 am to 5:00 pm**

**Yufeng Liu**
University of North Carolina at Chapel Hill
**Genevera Allen**
Rice University

**Overview:** This full day short course will provide an overview of statistical machine learning techniques with applications to the analysis of big biomedical data. Both supervised and unsupervised techniques will be covered. Supervised learning techniques include penalized regression such as LASSO and it variants, support vector machines, Boosting, and tree-based methods. Unsupervervised learning techniques include dimension reduction methods such as principal components analysis and non-negative matrix factorization, clustering analysis, and network analysis with graphical models. The main emphasis will be on the analysis of real high-dimensional data sets from various scientific fields, including genomics and biomedical imaging. The techniques discussed will be demonstrated in R.

**Instructors Biography:** Yufeng Liu is professor in Department of Statistics and Operations Research, Department of Biostatistics, and Department of

*Continued on following page....*

Genetics at UNC-Chapel Hill. His current research interests include statistical machine learning, high dimensional data analysis, and bioinformatics. He has taught statistical machine learning courses multiple times at UNC, as well as short courses on this subject at Joint Statistical Meetings and Biostatistics Summer Institutes at University of Washington. Dr. Liu received the CAREER Award from National Science Foundation in 2008, and Ruth and Phillip Hettleman Prize for Artistic and Scholarly Achievement in 2010. He is a fellow at American Statistical Association and an elected member of International Statistical Institute.

Genevera Allen is the Dobelman Family Junior Chair and Assistant Professor in Statistics and Electrical and Computer Engineering at Rice University. She also holds a joint appointment at the Jan and Dan Duncan Neurological Research Institute at Baylor College of Medicine. Dr. Allen's research focuses on developing statistical methods to help scientists make sense of their 'big data' in applications such as high-throughput genomics and neuroimaging. Her work lies in the areas of modern multivariate analysis, graphical models, statistical machine learning, and data integration or data fusion. She has taught statistical machine learning courses for many years at Rice University as well as short courses on this subject at the Biostatistics Summer Institutes at University of Washington and the Data Science Summer Institute at Rice University. Dr. Allen has represented the American Statistical Association in the "This is Statistics" public relations campaign and is the recipient of several honors including the David P. Byar Young Investigator travel award and the International Biometric Society's Young Statistician Showcase award. Current emerging area is genomics and modern genetics (GWAS and sequencing data).

## SC4: Statistical Genetic and Genomic Analysis of Complex Traits with GWAS and Next Generation Sequencing Data

**HALF DAY: 8:00 am to 12:00 noon**

**Hongyu Zhao**
Yale University
**Fei Zou**
University of North Carolina at Chapel Hill

**Overview**: Recent advances in genotyping and sequencing technologies have revolutionized biological and biomedical research. Great efforts have been taken to uncover genetic factors of many complex human disorders and traits with genome-wide association and next generation sequencing studies. An unprecedented wealth of data generated from these studies has created great opportunities and challenges for statisticians and biomedical researchers. In this short course, we will cover genotyping and sequencing technologies, and state-of-the-art statistical approaches to analyzing and interpreting GWAS and Next Generation Sequencing data. Topics to be covered include:

- Introduction to genetics, and genomic analysis technologies and software
- Introduction to GWAS data processing and analysis
- Introduction to next generation sequencing data and analysis
- Real data examples on psychiatric diseases, autoimmune disorders, cancer, and anthropological traits.

**Instructors Biography:** Dr. Hongyu Zhao is the Ira V. Hiscock Professor of Biostatistics and Professor of Statistics and Genetics, Chair of the Biostatistics Department and the Co-Director of Graduate Studies of the Inter-Departmental Program in Computational Biology and Bioinformatics at Yale University. His research interests are the applications of statistical methods in molecular biology, genetics, drug developments, and personalized medicine. Some of his recent projects include large scale genome wide studies to identify genetic variants underlying complex diseases (schizophrenia, bipolar, autism, and substance abuse), biological network modeling and analysis, disease biomarker identification through proteomics, genome annotations, microbiome analysis and systems biology study of herbal medicine. Dr. Zhao has been teaching statistical genetics and bioinformatics at Yale since 1996 and has offered multiple short courses on statistical genetics and genomics at JSM, ENAR, Deming Conference, and other statistical and genomics conferences.

Dr. Fei Zou is Professor in Department of Biostatistics and Department of Genetics at UNC-Chapel Hill. She is also the director of the Biostatistics/Bioinformatics core of UNC Neuroscience Center Research Cores. Her research interests include genome-wide association mapping with emphasis on assessing population stratification, and bias correction; QTL mapping with

*Continued on following page....*

experimental mouse data; integrated statistical analysis of high-dimensional genetics and genomics data, and differential gene expression and methylation analysis of next generation sequencing data. Dr. Zou is a co-investigator on multiple association studies on schizophrenia, early child brain development, and cystic fibrosis. She has taught statistical genetics and genomic courses multiple times at UNC and SAMSI.

## SC5: Bayesian Evidence Synthesis in Medicine

**HALF DAY: 8:00 am to 12:00 noon**

**Heinz Schmidli**
Novartis, Basel, Switzerland
**David Ohlssen**
Novartis, New Jersey

**Overview:** In the past decade, evidence synthesis has become an indispensable approach in many areas of medicine. Using combined information from several clinical studies, the effectiveness and safety of available treatments can be compared, providing guidance on how to best treat patients. Evidence synthesis approaches are also important in the design and analysis of clinical studies, for example when planning to use historical placebo information, or in non-inferiority studies. Bayesian methods are ideally suited to combine information from various sources. This course will describe the main meta-analytic models for evidence synthesis of clinical trials. These models take into account differences among studies by use of covariate information and random-effects. Applications and examples will motivate and illustrate the methods, covering topics such as comparative effectiveness research, clinical trials with historical controls, non-inferiority and biosimilar clinical trials, subgroup meta-analyses, and safety meta-analyses. The course will cover:

- Evidence synthesis in medicine
- Brief introduction to Bayesian inference and computation
- Evidence synthesis tools: Meta-analysis, Network meta-analysis, Meta-regression.
- Prediction based on meta-analytic models
- Assessing and dealing with conflicting information

**Instructors Biography:** Dr. Heinz Schmidli is a Biometrical Fellow within the Novartis Statistical Methodology group, based in Basel, Switzerland. He studied Mathematics at the University of Basel, and received his PhD in Statistics in 1994. In 2012 he received the Paul-Martini-Prize of the GMDS (jointly with Tim Friede), and in 2013 the Novartis leading scientist award. He has authored or co-authored more than 50 articles in peer-reviewed journals, and is author of a book.

Dr. David Ohlssen is currently a Biometrical Fellow and Bayesian focus team lead, within the Novartis statistical methodology group, based in East Hanover New Jersey. Since joining Novartis in 2007, he has developed a broad range of experience in applying novel statistical approaches within a drug development setting. Previously, after completing his PhD in Biostatistics at the University of Cambridge, he worked as a research fellow at the MRC Biostatistics Unit (Cambridge UK), where his interests included: diagnostics for Bayesian models, novel clinical trial design and statistical methods for the profiling of health-care providers. His professional activities include serving as a member of the Bayesian DIA Working Group and within the group acting as the chair of the safety meta-analysis sub-team.

## SC6: Practical Solutions for Simple Problems with Bad Consequences in Clinical Trials

**HALF DAY: 1:00 pm to 5:00 pm**

**Peter Thall**
University of Texas, M.D. Anderson Cancer Center

**Overview** : Many conventional methods used for clinical trial design or analysis have undesirable properties that are not obvious and often are not well understood. In many cases, the problem may have very undesirable consequences if it is ignored. This half day short course will describe a variety of such problems, and provide a practical alternative for each. Topics will include (1) consequences of ignoring treatment efficacy in dose-finding trials, (2) misinterpreting hypothesis test-based designs, (3) dealing with late onset toxicities, (4) futility or safety monitoring rules that may not work well in practice, (5) consequences of ignoring patient heterogeneity (6) scientific and ethical problems with adaptive randomization in comparative trials, (7) estimation bias, (8) some counterintuitive relationships between early response rate and mean survival time, (9) decision making for multi-stage dynamic treatment regimes, and (10) SMART (Sequential Multiple Assignment

*Continued on following page....*


**"Greetings From Austin - Capitol of Texas"** *postcard mural, at the Roadhouse Relics building*



Randomized Trial) designs. The examples will include very little mathematical detail, but attendees should have some knowledge of elementary probability and statistics. The course is appropriate for anyone involved in clinical trial design, conduct, or analysis, including statisticians, physicians, research nurses, administrators, professionals in the pharmaceutical industry, or federal employees in the NIH or FDA.

**Instructor Biography:** Peter Thall has been a Science Faculty member at M.D. Anderson Cancer Center since 1990, where he holds the Anise J. Sorrell Endowed Professorship. He is a Fellow of the American Statistical Association (ASA) and the Society of Clinical Trials, and received the Don Owen award from the San Antonio chapter of the ASA in 2014. Dr. Thall has pioneered application of Bayesian methods in medical research, designed hundreds of clinical trials, published over 200 papers and book chapters in the statistical and medical literature, and presented 28 short courses and over 190 invited talks. His current research interests include Bayesian utility-based clinical trial design, nonparametric Bayesian methods, bias correction, and dynamic treatment regimes. He is an Associate Editor for the journals Clinical Trials and Statistics in Biosciences, an ASA Media Expert, a member of the International Biometric Society ENAR Regional Advisory Board, and is Principal Investigator of the NIH/NCI RO1 grant "Statistical Methods for Complex Cancer Trials."

| | REGISTRATION RECEIVED | |
|---|---|---|
| | By Feb.1 | After Feb.1 |
| **Member** | $75 | $85 |
| **Non-Member** | $85 | $95 |
| **Student** | $40 | $50 |

## T1: Topics in High-Performance Computing with R

**Monday, March 7 | 8:30 am to 10:15 am**

**John W. Emerson**
Yale University

**Description:** This tutorial will introduce you to topics in high-performance computing with R. We will touch upon a few important language fundamentals relating to memory management and algorithmic efficiency. We will quickly explore the parallel package (containing snow and multicore), but will concentrate on the elegant framework for parallel programming offered by packages for each and the associated parallel backends. We will conclude with basic examples of handling larger-than-RAM numeric matrices and use of shared memory. Take-away material will include a concise example authoring an R package that includes C/C++ code, uses Rcpp, and roxygen2 for efficient documentation.

**Instructor Biography:** John W. Emerson (Jay) is Director of Graduate Studies in the Department of Statistics at Yale University. He teaches a range of graduate and undergraduate courses as well as workshops, tutorials, and short courses at all levels around the world. His interests are in computational statistics and graphics, and his applied work ranges from topics in sports statistics to bioinformatics, environmental statistics, and Big Data challenges. He is the author of several R packages including bcp (for Bayesian change point analysis), bigmemory and sister packages (towards a scalable solution for statistical computing with massive data), and gpairs (for generalized pairs plots). His teaching style is engaging and his workshops are active, hands-on learning experiences.

## T2: Interactive Data Visualizations in R with shiny and ggplot2

**Monday, March 7 | 10:30 am to 12:15 pm**

**Garrett Grolemund**
RStudio

**Description:** Data visualizations are one of the most useful tools for scientific discovery and communication. They become even more useful when they are enhanced with interactive techniques like linked brushing, hover effects, and zooming. However, interactive visualizations typically require specialized or proprietary software that cannot be easily installed or extended. As a result this valuable technique often goes underused. This tutorial will teach you an easy workflow for creating both static and interactive data visualizations with R, a free and open source computer language that is widely used by biostatisticians. You will learn how to make static plots with the grammar of graphics, an easy to use system for building plots. With the grammar of graphics, you can describe---and build---any plot by describing three components: a data set to visualize, a geometric object to use to represent individual observations, and a set of mappings between variables in the data set and visual properties of the geometric objects (like location, size, and color). We will use the popular ggplot2 package to implement the grammar of graphics. We will then look at how to add interactivity to your plots with Shiny, an R package that builds interactive data displays. We will focus on the newest features of Shiny, which enable interactive graphics. Participants will finish the tutorial by building their own visualizations that use linked brushing and hover effects.

**Instructor Biography:** Garrett Grolemund is a Data Scientist and Master Instructor at RStudio. He holds a

*Continued on following page....*

Ph.D. in Statistics and specializes in teaching others how to do data science with open source tools. He is a long time user and advocate of R; he wrote the popular lubridate package for working with dates and times in R, and is the editor of the Shiny development center at shiny.rstudio.com, which is the primary resource for learning how to build interactive web applications with R. Garrett designed and delivered the highly rated video series "Introduction to Data Science with R" by O'Reilly media, and he is the author of Hands-On Programming with R, as well as the co-author, with Hadley Wickham, of R for Data Science, a forthcoming book by O'Reilly media.

## T3: Clinical Trial Designs for Validating Prognostic and Predictive Markers in Oncology

**Monday, March 7 | 1:45 pm – 3:30 pm**

**Daniel J. Sargent**
Mayo Clinic

**Description**: Increasing scientific knowledge is creating both substantial opportunities and challenges in oncology drug develop. As diseases are sub-stratified into often biomarker-based groups, usual paradigms for phase II and III disease may no longer apply. In some circumstances, carefully conducted retrospective-prospective analysis may provide sufficient evidence of a predictive biomarker for clinical use. Prospectively, enrichment designs are appropriate when preliminary evidence suggest that patients with/without that marker profile do not benefit from treatments in question; however this may leave questions unanswered regarding the activity of an agent in a larger but still relevant population. An unselected design is optimal where preliminary evidence regarding treatment benefit and assay reproducibility is uncertain. The biomarker-based strategy design may be useful when there is a choice between many treatment options. Adaptive analysis designs allow for pre-specified marker defined subgroup analyses of data from a RCT. Umbrella or basket trials enroll large groups of patients with subsequent assignment to either individual randomized trials or single arm investigations. These trials may be disease specific, or may include patients from multiple sites who share a common biomarker status. We discuss features of these various novel design strategies in the context of multiple ongoing and planned real trials. Emphasis will be placed on practical considerations that may impact an academically optimal design.

**Instructor Biography:** Daniel J. Sargent, Ph.D. is the Ralph S. and Beverly E Caulkins Professor of Cancer Research at the Mayo Clinic. He is the Group Statistician for the Alliance for Clinical Trials in Oncology and the Director of Biostatistics Shared Resource at the Mayo Clinic Comprehensive Cancer Center. Dr. Sargent co-chaired a joint NCI-EORTC committee on methodology for tumor marker studies, was a member of the FDA panel on endpoints for colon cancer clinical trials, and from 2007 - 2013 was a member of the US NCI Clinical Trials Advisory Committee, which oversees all NSI funded cancer clinical trials in the United States. He presently serves on the Clinical Trial Design Task force of the NCI's Investigational Drug Steering Committee. He has published extensively in colorectal cancer treatment in the elderly, optimal clinical trial endpoints, and prognostic and predictive biomarkers. He has authored over 290 peer-reviewed manuscripts, book chapters, editorials, and letters.

## T4: A New Paradigm for Finding the Subset of Patients who Benefit from a Test Treatment

**Monday, March 7 | 3:45 pm – 5:30 pm**

**Richard Simon**
National Cancer Institute
**Noah Simon**
University of Washington

**Description**: Conventionally defined diseases are often heterogeneous in biology, natural course and response to treatment. The traditional paradigm of conducting broad eligibility randomized clinical trials followed by post-hoc subset analysis does not provide a reliable or efficient basis for precision/personalized medicine. Broad eligibility clinical trials often lead to subsequent over-treatment of patient populations and conventional post-hoc subset analysis often does not provide reliable or actionable findings. In this tutorial we describe an alternative paradigm for finding and evaluating the subset of patients who benefit from a test treatment relative to a control. The paradigm re-formulates the problem as one of developing a predictive classifier and obtaining valid estimates of the properties of the classifier such as the treatment effect in the population predicted to benefit from the test treatment. We describe two contexts for the application of this paradigm. One context involves prospective

*Continued on following page....*

application using an Adaptive Enrichment Design. The other context involves retrospective application using the principles of the Prospective/Retrospective Design. We describe the development of predictive classifiers for binary outcome and time-to-event data. Several types of predictive classifiers are described. The approaches are applicable to settings with a small number of candidate features are available and settings with high-dimensional feature sets. We describe how the approach provides an intended use population and describe use of re-sampling and permutation techniques to obtain valid estimates of treatment effect for that population. Recent results by us and related work by others will be described.

**Instructors Biography:** Richard Simon is Chief of the Biometric Research Program at the National Cancer Institute and head of the Computational and Systems Oncology Branch. He is a fellow of the American Statistical Association and a former member of the Oncologic Drug Advisory Committee of the FDA. He is the architect of BRB Array Tools software and author of Using Genomics in Clinical Trials and Predictive Medicine (Cambridge U. Press 2013). He is the recipient of the 2013 Karl Peace award of the American Statistical Association "for contributions that have played a pivotal role in bridging the gap among statistics, clinical research, and translational medicine to improve human health".

Noah Simon (Ph.D. in Statistics from Stanford University) is an Assistant Professor in the Department of Biostatistics at the University of Washington. He works on problems at the intersection of statistics, biology and computer science, developing algorithms to build scientific knowledge from modern high-throughput technologies. His interests include high dimensional modeling and inference, selection-bias in high-throughput experiments, biomarker development, and adaptive clinical trial design. He was a Weiland Fellow and currently holds the Genentech Endowed Professorship in Biostatistics at the University of Washington

## T5: Introduction to High Throughput DNA Sequence Data Analysis Using R / Bioconductor

**Tuesday, March 8 | 8:30 am – 10:15 am**

**Martin Morgan**
Roswell Park Cancer Institute

**Description**: Modern methods of high-throughput genomic data generate large primary data sets that require significant data manipulation and statistical summary before arriving at biological insight. This workshop starts by outlining basic DNA sequence analysis work flows, from primary data generation to biological interpretation. We use this outline, and especially the 'RNA-seq known gene differential expression' work flow, to identify relevant data management and statistical issues. The workshop then steps through R and Bioconductor code to implement essential stages in data management and statistical analysis. We conclude by briefly contrasting differences in the biological, technological and statistical aspects of RNA-, DNA-, and methyl-seq, with a brief overview of the resources available for further study.

**Instructor Biography:** Dr. Martin Morgan leads the successful open source, open development Bioconductor project (http://bioconductor.org) for the analysis and comprehension of high throughput genomic data. Dr. Morgan's interests include statistical computation, integrative analysis of multiple 'omics data sets, and effective data comprehension.

## T6: Adaptive designs for Confirmatory Clinical Trials

**Tuesday, March 8 | 1:45 pm – 3:30 pm**

**Franz König**
Medical University of Vienna

**Description:** Since the first methodological papers on adaptive designs, some published more than 25 years ago, adaptive designs have gained increasing attention in drug development. Especially in pivotal phase III trials, their use is subject to enhanced scrutiny by regulators as the increased complexity of flexible study designs also increases the risk of operational and statistical biases and hidden fallacies. Broad enthusiasm about potential applications of such designs faced critical positions regarding their statistical efficiency. Despite, or possibly because of, this controversy, the methodology and its areas of applications grew steadily over the years, with significant contributions from statisticians working in academia, industry and agencies around the world. In the meantime, such types of adaptive designs have become the subject of three major regulatory guidance documents in the US and Europe and the field is still evolving. The main goal of this tutorial is to give an introduction to the key principles and statistical methodologies of adaptive designs for confirmatory clinical trials. Important applications of

*Continued on following page....*

adaptive designs include sample size reassessment, treatment selection procedures, and population enrichment designs. The change of design parameters at an adaptive interim analysis may depend on any internal and external data available. Using adaptive multiple test procedures the type I error rate can be controlled even if the selection rule, the number of selected treatments or the final sample sizes are not prefixed. The tutorial shall provide an overview of methods from the published literature including the most recent developments. Special emphasis is put on sample size reassessment and multiple hypotheses testing with adaptive designs. Regulatory issues and case studies will be discussed.

**Instructor Biography**: Franz König is an Associate Professor at the Section of Medical Statistics at the Medical University of Vienna, Austria. He serves on several data safety monitoring boards (DSMB) as independent statistical expert. From 2008 till 2010 he was seconded to the European Medicines Agency (London, UK) as statistical expert in the Unit Human Medicines Development and Evaluation. At the EMA he held the Scientific Secretariat of the then newly founded Biostatistics Working Party (BSWP). He was involved in the development of guidelines and assessment of statistical methods and clinical trial protocols. His main research interests are multiple testing, adaptive/flexible designs, interim analyses and data safety monitoring boards. Franz has served as Guest Editor for Special Issues in Biometrical Journal and Statistics in Medicine. He is currently the work package leader of the work package "adaptive designs" in the EU FP7-funded research project IDEAL and deputy coordinator of an EU Horizon 2020 funded Marie Curie ITN network IDEAS on early drug development studies.

## T7: Quantile Regression for Survival Analysis

**Tuesday, March 8 | 3:45 pm – 5:30 pm**

**Limin Peng**
Emory University

**Description:** Quantile regression offers a useful alternative strategy for analyzing survival data. It formulates covariate effects on the quantile(s) of an event time of interest. Such a modeling perspective entails easy interpretations of covariate effects, and allows for a comprehensive and flexible evaluation of the association between covariates and the survival outcome. Moreover, many existing quantile regression methods for survival analysis enjoy simple and stable computation.
By these nice features, quantile regression has emerged as a valuable practical tool that can provide in-depth investigations of survival studies. In this tutorial, I plan to introduce the fundamentals of quantile regression (including basic concepts, modeling strategies, computing features and interpretations). Through reviewing and comparing with traditional regression methods for survival data, I will elaborate the benefits of conducting quantile regression analysis in survival settings. I will provide an overview of currently available methods for quantile regression with survival data, with a particular focus on a few popular approaches developed for randomly censored data. Available computing resources will be discussed. Examples will be presented to illustrate method applications and implementation.

**Instructor Biography:** Dr. Limin Peng is an Associate Professor in the Department of Biostatistics and Bioinformatics at the Emory University. She has been an active researcher in the areas of survival analysis and quantile regression. She has published a series of work in major statistical journals on quantile regression methods (or extensions) for various types of survival data, including randomly censored data, competing risks data, semi-competing risks data, truncated data, censored longitudinal data, and recurrent events data. Dr. Peng has disseminated novel applications of quantile regressions to research communities in Neurology, Dialysis, and Cystic Fibrosis. Dr. Peng has extensive teaching experiences on topics related to this tutorial. She has taught courses, "Survival Analysis Methods" and "Quantile Regression" at Emory University and taught a tutorial on quantile regression for survival analysis at 2014 ENAR Spring meeting and a short course on the same topic at Centers for Disease Control and Prevention.

# ENAR 2016 Round Tables

## Registration is Required **$40.00**

Monday March 7 | 12:15 pm - 1:30 pm

### R1: Challenges in Cancer Epidemiology in the Era of Genomic Medicine

**Colin Begg**
Sloan Kettering Cancer Center

**Description:** For decades cancer epidemiologists have investigated cancer risk associations using studies in which the disease has been studied with the anatomic site of origin of the cancer as the organizing framework. In the clinical research arena the advent of powerful technologies for examining tumors at the molecular level are altering the landscape of cancer investigation to one in which cancers are increasingly being classified on the basis of somatic mutations and other tumor characteristics. In this session we will discuss the implications of these trends for the design and analysis of epidemiologic studies of cancer risk.

SOLD OUT

### R2: Statistical Issues in the Analysis of Electronic Health Record Data

**Sebatstein Haneuse**
Harvard University

**Description:** Electronic health record (EHR) data present an incredibly appealing opportunity for public health and medical research. Relative to data from dedicated designs, EHR data often contain rich information on a broad range of patients over long timeframes and in real time. Furthermore, they are typically relatively cheap to obtain. Nevertheless, the use of EHR data for research purposes is subject to many challenges including: (i) accurate extraction of text-based information; (ii) missing data; (iii) measurement error and misclassification and (iv) the control of confounding. While these issues arise in other contexts, the complexity and high-dimensional nature of EHR data poses additional challenges for statisticians. The purpose of this roundtable is to provide a forum for statisticians and analysts to discuss these issues, in general as well as in their own contexts, and also to discuss recent methodologic developments in the area.

SOLD OUT

### R3: Precision Medicine

**Michael Kosorok**
University of North Carolina at Chapel Hill

**Description:** We will discuss recent progress in precision medicine and key roles that biostatistics plays, including important aspects of both discovery and evaluation. We will also examine the multi-disciplinary characteristics of this research and how recent developments in statistical machine learning are being driven by biostatistical considerations.

SOLD OUT

### R4: Big Data in Healthcare Evaluation: The Role for Biostatistics

**David Ohlssen**
Novartis

**Description:** Traditionally, big data has been defined as a dataset too large to be efficiently collected, stored, organized, and analyzed using currently available software. This complexity can be characterized by a combination of data volume, velocity and variety. It has been argued that the use of big data has enormous potential in the healthcare sector. For example, the aggregation and analysis of large real-world datasets adds power, allows trends and associations to be more readily observed, and overcomes some of the limitations associated with randomized clinical trials. The growing area of big data provides enormous potential for contributions by biostatisticians. However, so far our profession and ENAR have not been highly involved in big data activities. Within the context of healthcare evaluation this roundtable aims to discuss big data and the role for biostatistics. Possible topics for discussion include:

1. The skills we need to be involved with big data
2. Combining scientific knowledge (e.g. knowledge about pharmacology or design) with big data
3. Is big data just for prediction or can we make progress with causal inference

SOLD OUT

*Continued on following page....*

4. Trials of the future will utilize censors and gadgets that potentially provide much richer more complex data from clinical research
5. Data transparency initiatives have led to most pharmaceutical companies providing access to their individual patient clinical trial. What can be done to explore this new abundance of clinical data
6. Computation developments such as Hamilton Monte Carlo and STAN

## R5: Quantitative Safety Review at FDA's Center for Drug Evaluation and Research (CDER)

**Mark Levenson**
Food and Drug Administration

**Description:** This round table will discuss recent examples and future directions of quantitative safety review at FDA's Center for Drug Evaluation and Research (CDER). The following topics will be the bases of the discussion:

SOLD OUT

» Examples of quantitative safety review including meta-analyses, large safety trials, and observational studies conducted by drug companies or the FDA to address important drug safety issues.
» Implications of recent and proposed legislation and guidances and new data sources.
» Statistical careers at the FDA/CDER, particularly those involving drug safety.

## R6: Leadership in Statistics

**Dubois Bowman**
Columbia University

**Description:** Statistical leadership is critical given the growing demand for analytical skills in our society, coupled with the emergence of contributions from varied quantitative disciplines. In this roundtable, we will discuss leadership opportunities for statisticians. We will examine key elements of being an effective leader, both within one's field and in interdisciplinary settings. We will also cover ways that statisticians can prepare for leadership roles during various stages of their careers and tips for identifying leadership positions.

SOLD OUT

## R7: Balancing Responsibilities in Academia: Methodological and Collaborative Research, Teaching, and Service

**Amy Herring**
University of North Carolina at Chapel Hill

**Description:** Faculty positions in biostatistics provide many interesting opportunities to engage in methodological and collaborative health science research, collaborate with students and junior scholars, and improve the home institution and profession more broadly through service activities. During this roundtable, we will share goals and formulate strategies for success (and recipes for stress) in exploring the sample space of academic opportunities. Amy Herring is Professor and Associate Chair of Biostatistics at UNC-Chapel Hill. She is a former ENAR President and recently completed a term as Chair of UNC's Committee on Appointment, Promotion, and Tenure.

## R8: Publishing Without Perishing: Strategies for Success in Publishing in Biostatistical Journals

**Marie Davidian**
North Carolina State University

**Description**: Contributing to the advance of your discipline through publication of articles in peer-reviewed journals is a fundamental expectation for both junior and not-so-junior biostatistical researchers alike. Success in publishing one's work ensures that it will be widely disseminated to researchers and practitioners who stand to benefit. In addition, funding agencies and academic institutions place considerable importance on a successful record of publication. Accordingly, understanding the peer review and editorial processes of top journals and mastering the art of writing an effective journal article are keys to success in publishing. How does one determine the best outlet for one's work? What are the essential elements of a successful journal article? How does one maximize the chance of acceptance? What strategies can ensure that a published paper is read and cited? How does one make optimal use of limited space and additional supplementary material in conveying the message? What are the roles of the editor, associate editor, and referees? What considerations do editors use when evaluating a paper? This roundtable will provide a forum for candid discussion of these and other questions.

SOLD OUT

## R9: Practical Considerations for Teaching Biostatistics in a Hybrid, Blended, or Online Format

**Jane Monaco**
University of North Carolina at Chapel Hill
**Todd Schwartz**
University of North Carolina at Chapel Hill

SOLD OUT

**Description:** Whether you are considering updating an existing biostatistics course or are tasked with developing a new one, numerous options exist for delivering the content. Many courses are moving away from traditional classroom lectures. While some courses are taught completely online, debate surrounding the effectiveness of this format continues. Meanwhile, gaining in popularity are "Hybrid" or "Blended" formats, in which students are expected to view pre-recorded lectures before coming to class; class-time is devoted to active learning strategies and hands-on practice. This roundtable will provide an opportunity to discuss practical considerations, best practices and lessons learned in teaching biostatistics as classroom technology, online resources and student expectations continue to evolve. Discussion topics will include:

- What technologies are most effective? Considerations include ease of use, compatibility, and cost. Technology to be discussed may include desktop recording software, grading options, and classroom technology.
- What resources are available? Rather than "reinventing the wheel," what existing options are helpful and available online?
- How can you avoid a "flipped classroom" flop? What strategies are effective for delivering content and administering the course?
- Roadblocks – what happens when students or administrators are reluctant to embrace an innovative classroom model?
- What is the time commitment for developing or updating a course with a hybrid, blended or online format?



### Participate in Student-Focused Elements of the Scientific Program

The Sunday night mixer presents an ideal opportunity to obtain feedback on your work in our Annual ENAR Poster session. This year we will conduct our fifth Poster Competition for the session. Prizes will be announced within topical areas in the Tuesday morning Presidential Invited Address session. A student winner will be selected within each topical area. Watch for details on entering the competition on the website when the meeting registration becomes available.

### Join Us for the Tuesday Evening Dinner and Social Event

SOLD OUT

Reduced registration fee offered to students to attend (see page 7).

### Educational and Professional Development Opportunities

Be sure to take advantage of the educational offerings to be held during the meeting – short courses, tutorials, and roundtable discussions (see pages 94 - 105).

### Network with Your Fellow Students

Back by popular demand, a CENS mixer will be held the evening of Monday, March 7, 2016. This is a great way to meet and greet your students from other graduate programs. Don't miss this opportunity to begin building connections with your future colleagues and friends.