



# ENAR 2019 SPRING MEETING

WITH IMS & SECTIONS OF ASA  
MARCH 24-27, 2019  
MARRIOTT PHILADELPHIA, PHILADELPHIA, PA

## PROGRAM











# TABLE OF CONTENTS

- 4 [Welcome and Overview](#)
- 6 [Acknowledgements](#)
- 9 [Special Thanks](#)
- 10 [Philadelphia Highlights](#)
- 12 [Presidential Invited Speaker](#)
- 13 [Program Summary](#)
- 18 [Scientific Program](#)
- 60 [Short Courses](#)
- 62 [Tutorials](#)
- 63 [Roundtables](#)
- 64 [CENS Mission](#)
- 66 [Floor Plan](#)



ENAR 2019 Spring Meeting  
With IMS & Sections of ASA  
March 24-27 | Marriott Philadelphia, Philadelphia, PA

# WELCOME

I am delighted to welcome you to the 2019 ENAR Spring Meeting! I extend a special welcome to those attending their first ENAR and hope you will enjoy the Spring Meeting and return for many future meetings plus get more involved in our organization!

The 2019 ENAR Spring Meeting will be held at the Philadelphia Marriott Downtown. Philadelphia is home to Independence Hall, the Liberty Bell, National Constitution Center, Museum of the American Revolution, Philadelphia Museum of Art, and many more superb museums and attractions. There are lots of great restaurants in the area, plus the Reading Terminal Market across the street for a quick bite.

The four-day meeting, March 24-27, 2019, will host students, researchers, and practitioners from all over our biostatistics profession, from academia to industry and government, from places large and small, brought together to share ideas, learn and connect over our joint interest in biometry. Continuing in ENAR's strong tradition, the meeting will offer numerous opportunities throughout the Scientific and Educational programs to discover the latest developments in the field, learn about state-of-the-art research methods and software, see how statistics can inform policy and decision-making, and hear about novel applications of statistical methods to many areas in the life sciences. Additionally, the meetings are a fantastic opportunity for professional networking, meeting new people, connecting job seekers with employers, and reconnecting with friends and colleagues. There will be opportunities to check out the latest textbooks and see demonstrations of new software from our exhibitors and vendors, who have partnered with ENAR.

As a professional organization composed of a diverse group of individual members, ENAR is committed to fostering a culture of inclusion, professionalism and civil discourse that cultivates an environment where ideas are exchanged openly and freely with mutual respect and trust. ENAR has adopted a Meeting Conduct Policy intended to guide all attendees at ENAR's annual Spring Meeting and attendees will be required to assent to the policy as part of registration going forward.

The ENAR Spring Meeting is only possible through the efforts of a large number of hard-working volunteers. Without their time, energy and ideas, it would be impossible to coordinate and organize the program and manage the meeting logistics. I am sincerely thankful for each one of you! Your hard work and commitment are critical to the success of the meeting!

## Scientific Program

Through the leadership of the Program Chair Pamela Shaw (University of Pennsylvania) and Associate Chair Michael Fay (NIAID), and contributions from many of you, the Program Committee (consisting of 10 ASA section representatives and 4 at-large ENAR members) has assembled a diverse and exciting invited program. The sessions cover a wide range of topics, including statistical advances for microbiome data, electronic health records data, wearable/mobile technology, self-reported outcomes, non-ignorable missing data, data integration, causal inference, survival outcomes, spatial modeling, precision medicine, and clinical trials. The IMS Program Chair Vladimir Minin (University of California, Irvine) has also put together complementary sessions on classification, variable selection, causal inference, statistical modeling in cell biology, microbiome data, surveillance data and mediation analysis for high-dimensional data.

Poster sessions play a prominent role at the ENAR Spring Meeting, and continue to be a vital part of the program. In addition to contributed and invited posters, the 2019 ENAR Spring Meeting will continue contributed SPEED poster sessions, in which presenters give a two-minute elevator speech on the highlights of their posters. Unlike previous years, the speed sessions will utilize digital poster boards in 2019, giving presenters the opportunity for more interactive posters. Monday, March 25th

will feature the thematically grouped contributed speed poster sessions. These will feature two invited posters from well-known researchers and will run parallel with the rest of the sessions in the scientific program. As in previous years, the regular contributed posters will be featured during the Opening Mixer on Sunday evening. This year, poster presenters will be assigned one-hour slots to be available at their poster, giving everyone a chance to view the amazing research on display. Posters in this session will be eligible to win an award as part of the popular ENAR Regional Advisory Board's poster competition!

## Educational Program

Our educational program provides lots of opportunities to learn a new area of statistical techniques, develop new computational skills, and to discuss the latest research or career development skills with some leading experts. The Educational Advisory Committee has assembled an enriching suite of short courses, tutorials and roundtables covering a wide range of topics from renowned instructors.

Short course topics include clinical trials design using historical data, design of observational matched studies, big data and data science, analysis of medical cost data, subgroup identification, and StatTag for reproducibility inside Word documents, and smart simulations in SAS and R. Tutorial topics include an introduction to causal effect estimation, python programming, modern multiple imputation, meta-analysis of clinical trials and data visualizations with ggplot2. Roundtable luncheons provide a more intimate discussion with distinguished statisticians from academia, government and industry. Topics range from Time Management: Taming the Inbox to Creating a Research Group across Campus/Regionally and analytic challenges in clinical trials or administrative health data. Be sure to take a look and sign up for something interesting!

I would like to extend a special thanks to the members of the Educational Advisory Committee - Ofer Harel (University of Connecticut), Zhen Chen (NICHD), Guofen Yan (University of Virginia), and Devan Mehrotra (Merck) - for their insights and efforts in putting together such an outstanding educational program.

## Keynote Lecture

I am thrilled to announce that the 2019 ENAR Presidential Invited Address will be given by Dr. Francesca Dominici, the Clarence James Gamble Professor of Biostatistics, Population and Data Science at the Harvard T.H. Chan School of Public Health and Co-Director of the Harvard Data Science Initiative. Dr. Dominici is a statistician and data scientist whose pioneering scientific contributions have advanced public health research around the globe. Her life's work has focused broadly on developing and advancing methods for the analysis of large, heterogeneous data sets to identify and understand the health impacts of environmental threats and inform policy. In 2015, she was awarded the Florence Nightingale David award based on her contributions as a role model to women and her demonstrated excellence in statistical research, leadership of multidisciplinary collaborative groups, statistics education and service to the profession of statistics. To learn more about Dr. Dominici and her Invited Address, please see page 11.

## Additional Meeting Activities

The ENAR 2019 Spring Meeting will feature a number of other activities in addition to the scientific and educational programs. Immediately preceding the Spring Meetings, a workshop for Junior Biostatisticians in Health Research will be held on Friday, March 22nd and Saturday, March 23rd. Organized by Drs. Howard Chang (Emory University), Betsy Ogburn (Johns Hopkins University), Jessica Franklin (Brigham and





Women's Hospital), David Vock (University of Minnesota), and Chris Slaughter (Vanderbilt University), this workshop aims to promote career development of junior investigators by bringing them together with a prestigious panel of senior investigators. Interested researchers should make sure they get their application in before the deadline.

On Sunday, March 24th, there will be the Fostering Diversity in Biostatistics Workshop, organized by Drs. Portia Exum (SAS Institute Inc.) and Felicia Simpson (Winston-Salem State University). Dr. Carmen Tekwe (Texas A&M School of Public Health) will serve as this year's keynote speaker. This workshop has been very popular and impactful and registration typically fills up quickly. Please be sure to register early if you are interested in attending!

Students, recent graduates, and other young professionals should plan to attend the Networking Mixer on Monday evening and the Tuesday luncheon event organized by the Council for Emerging and New Statisticians (CENS). These are fantastic opportunities for "younger" members to meet new people, learn about CENS and become more engaged with ENAR. Attendees seeking employment and prospective employers have the opportunity to connect via the Career Placement Center.



Tuesday evening will feature our first ENAR Sponsor and Exhibitor Mixer. This will be a great opportunity to check out the latest books and software. Please plan on joining the sponsors and exhibitors for the reception in the exhibition area after the last session on Tuesday. The evenings are also for catching up with friends, collaborators, and colleagues, and enjoying some of the wonderful sites, activities and delicious dining options nearby. The Local Arrangements Committee, chaired by Nandita Mitra (University of Pennsylvania), will provide some specific recommendations for attendees.

On a sadder note, the ENAR 2019 Spring Meeting will mark the last meeting that Kathy Hoskins (ENAR Executive Director) will be overseeing. Kathy will be retiring after her 23rd Spring Meeting! She will be leaving ENAR in the capable hands of Katie Earley. It has been a joy to work with Kathy and we wish her all the best in retirement!

**We hope to see you in Philadelphia for the 2019 ENAR Spring Meeting!**

**Sarah J. Ratcliffe**, ENAR 2019 President

**Kathy Hoskins**, ENAR Executive Director

**Katie Earley**, ENAR Deputy Executive Director



# ACKNOWLEDGEMENTS

ENAR would like to acknowledge the generous support of the 2019 Local Arrangements Committee, chaired by Nandita Mitra, and our student volunteers.

## We gratefully acknowledge NIH, and in particular the

National Cancer Institute  
National Heart, Lung, & Blood Institute  
National Institute of Environmental Health Sciences  
National Institute of Allergy and Infectious Diseases  
*For their generous support of the ENAR Junior Researchers Workshop*

## ENAR Coalition for Development of Junior Scholars

Drexel University  
Emory University  
ENAR  
Icahn School of Medicine at Mount Sinai  
Johns Hopkins Department of Biostatistics  
The Ohio State University  
The University of Rochester, Department of Biostatistics & Computational Biology  
Thomas Jefferson University

We gratefully acknowledge the invaluable support and generosity of our Sponsors and Exhibitors.

### Sponsors

AbbVie  
Cytel  
Eli Lilly and Company  
Emmes  
GlaxoSmithKline  
Janssen Research & Development  
Novartis  
SAS Institute Inc.  
StataCorp LLC  
Statistics Collaborative, Inc.

### Exhibitors

Cambridge University Press  
CENS – Council for Emerging and New Statisticians  
CRC Press, Taylor & Francis Group  
Flatiron HealthOxford University Press  
Penfield Search Partners  
SAS Institute Inc.  
Springer Science & Business Media LLC  
StataCorp LLC  
The Lotus Group LLC  
University of Kansas Department of Biostatistics  
Vanderbilt University Medical Center  
Wiley

## Executive Committee - Officers

President	Sarah Ratcliffe
Past President	Jeffrey Morris
President-Elect	Michael Daniels
Secretary (2017-2018)	Brisá Sanchez
Treasurer (2016-2017)	Reneé Moore

## Regional Committee (RECOM)

**President (Chair)** Sarah Ratcliffe  
**Nine ordinary members (elected to 3-year terms):** + Leslie McClure (RAB Chair)

2017-2019	2018-2020	2019-2021
Qi Long	Babette Brumback	Lynn Eberly
Simone Gray	Dean Follmann	Peter Song
Sean Simpson	Laura Hatfield	Alisa Stephens-Shields

## Regional Members Of The International Biometric Society Executive Board

Karen Bandeen-Roche, Joel Greenhouse, and José Pinheiro

## Regional Members Of The Council Of The International Biometric Society

KyungMann Kim, Nandita Mitra, Dionne Price and Brisa Sánchez

## Appointed Members Of Regional Advisory Board (3-Year Terms)

**Chair:** Leslie McClure  
**Secretary:** Emma Benn

2017-2019	2018-2020	2019-2021
Emma Benn	Naomi Brownstein	Ashley Buchanan
Manisha Desai	Jan Hannig	Emily Butler
Jeff Goldsmith	Eric Lock	Ani Eloyan
Elizabeth Handorf	Mark Meyer	Tanya Garcia
Cyrus Hoseyni	Taki Shinohara	Joseph Kang
Greg Levin	Daniela Sotres-Alvarez	Sung Duk Kim
Jiaqi Li	Peng Wei	Benjamin Risk
Yan Ma	Zhenke Wu	Ana-Maria Staicu
Therri Usher	Lin Zhang	Sameera Wijayawardana
	Lili Zhao	Yize Zhao

## CENS – Council for Emerging and New Statisticians

RAB Liaison: Elizabeth Hansdorf, Fox Chase Cancer Center  
Kylie Ainslie, Emory University  
Donald Hebert, The Emmes Corporation  
Shelley Liu, Icahn School of Medicine at Mount Sinai  
Alessandra Valcarcel, University of Pennsylvania  
Alexander Kaizer, University of Colorado-Anschutz Medical Campus  
Jeremiah Perez, Boston Biomedical Associates  
Fan Li, Duke University  
Jing Li, Indiana University Richard M. Fairbanks School of Public Health  
Will Eagan, Purdue University  
Nancy Murray, Emory University

## Programs

### 2019 Spring Meeting – Philadelphia

Program Chair: Pamela Shaw

Program Associate Chair: Michael Fay

### 2020 Spring Meeting - Nashville

Program Chair: Juned Siddique

Associate Program Chair: Chenguang Wang

### 2019 Joint Statistical Meeting

Michael Rosenblum

### 2020 Joint Statistical Meeting

TBD

### Biometrics Executive Editor

Geert Molenberghs

### Biometrics Co-Editors

Malka Gorfine

Stijn Vansteelandt

Debashis Ghosh

### Biometric Bulletin Editor

Havi Murad

### JABES Editor

Stephen T. Buckland

### ENAR Correspondent for the Biometric Bulletin

Jarcy Zee

### ENAR Executive Director

Kathy Hoskins

### International Biometric Society Executive Director

Peter Doherty

## Representatives

### Committee Of Presidents Of Statistical Societies (COPSS)

#### ENAR Representatives

Sarah Ratcliffe (President)

Michael Daniels (President-Elect)

Jeffrey Morris (Past President)

#### ENAR Standing/Continuing Committees

##### Nominating Committee (2018)

Chair: Scarlett Bellamy, Drexel University

Michael Daniels, University of Florida

Amy Herring, Duke University

Brian Millen, Eli Lilly and Company

Ruth Pfeiffer, National Cancer Institute/National Institutes of Health

Mary Sammel, University of Pennsylvania

Jeff Goldsmith, Columbia University

##### ENAR Webinar Committee (2018)

Michael Hudgens, University of North Carolina at Chapel Hill

Cory Zigler, Harvard School of Public Health

#### 2019 Junior Biostatisticians in Health Research Workshop

Howard Chang, Emory University

Betsy Ogburn, Johns Hopkins University

Jessica Franklin, Brigham and Women's Hospital

David Vock, University of Minnesota

Chris Slaughter, Vanderbilt University

with special thanks to Brisa Sánchez (University of Michigan) and

Brent Coull (Harvard University)

#### 2019 Fostering Diversity in Biostatistics Workshop

Portia Exum (Co-Chair), BIRD Business Intelligence Clients Testing Department, SAS Institute Inc

Felicia R. Simpson (Co-Chair), Department of Mathematics, Winston-Salem State University

Loni Philip Tabb, Department of Epidemiology and Biostatistics, Drexel University

Justine Herrera, Columbia University

Emma Benn, Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai

Vladimir Geneus, Eli Lilly & Company

Reneé H. Moore, Emory University, Rollins School of Public Health

Knashawn H. Morales, University of Pennsylvania, Perelman School of Medicine

Scarlett Bellamy, Department of Epidemiology and Biostatistics, Drexel University



# ACKNOWLEDGEMENTS

## Distinguished Student Paper Awards Committee

**Chair:** Scarlett Bellamy, Drexel University  
Shuo Chen, University of Maryland  
Haitao Chu, University of Minnesota  
Hakmook Kang, Vanderbilt University School of Medicine  
Ching-Ti Liu, Boston University  
Christine Mauro, Columbia University  
Loni Tabb, Drexel University  
Fang Chen, SAS Institute  
Jian Kang, University of Michigan  
Hernando Ombao, University of California at Irvine  
Hong Tian, Janssen R&D  
Xiaofei Wang, Duke University  
Dipankar Bandyopadhyay, Virginia Commonwealth University  
Ying Guo, Emory University  
Peisong Han, University of Waterloo  
Wenbin Lu, North Carolina State University  
Cory Zigler, Harvard University  
Fei Zou, University of North Carolina at Chapel Hill

## Student Award Winners

**Van Ryzin Award Winner**  
Xin Qiu, Janssen Research & Development

## Distinguished Student Paper Award Winners

Tramback Banerjee, University of Southern California  
Magdalena Bennett, Columbia University  
Roland Brown, University of Minnesota  
Katrina Devick, Harvard T.H. Chan School of Public Health  
Byung-Jun Kim, Virginia Tech  
Yujia Li, University of Pittsburgh  
Jeremiah Liu, Harvard University  
Lan Luo, University of Michigan  
Glen McGee, Harvard University  
Giovanni Nattino, The Ohio State University  
Tao Sun, University of Pittsburgh  
Zhe Sun, University of Pittsburgh  
Lili Wang, University of Michigan  
Theodore Westling, University of Pennsylvania  
Qiwei Wu, University of Missouri, Columbia  
Chong Wu, Florida State University  
Zhixing Xu, Florida State University  
Yishu Xue, University of Connecticut  
Yuan Yang, University of Michigan

Please visit the **ENAR website** ([www.enar.org](http://www.enar.org)) as a resource of information on all ENAR activities.





# SPECIAL THANKS

## Program Chair

Pamela A. Shaw, University of Pennsylvania

## Associate Program Chair

Michael P. Fay, NIAID

## IMS Program Chair

Vladimir N. Minin, University of California, Irvine

## Digital Program Coordinator

Alessandra M. Valcarcel, University of Pennsylvania

## Local Arrangements Chair

Nandita Mitra, University of Pennsylvania

## ASA Section Representatives

Haim Bar, University of Connecticut,  
ASA Health Policy Statistics Section

Inna Chevoneva, Thomas Jefferson University,  
ASA Statistics in Imaging Section

Yates Coley, Kaiser Permanente Washington Health Research Institute,  
ASA Biometrics Section

Meg Gamalo-Siebers, Eli Lilly,  
ASA Biopharmaceutical Section

Irina Gaynanova, Texas A&M University,  
ASA Statistical Learning & Data Mining Section

Samiran Ghosh, Wayne State University,  
ASA Mental Health Section

Charles B. Hall, Albert Einstein College of Medicine,  
ASA Statistics in Epidemiology Section

Yuan Huang, University of Iowa,  
ASA Statistics in Genomics & Genetics Section

Keyla Pagan-Rivera, Institute for Defense Analyses,  
ASA Statistics in Defense & National Security Section

Erin M. Schliep, University of Missouri,  
ASA Statistics & the Environment Section

## ENAR At-Large Members

Jian Kang, University of Michigan

Huilin Li, New York University

Sherri Rose, Harvard Medical School

Howard Chang, Emory University

## Educational Advisory Committee

Ofer Harel, University of Connecticut

Zhen Chen, National Institutes of Health

Guofen Yan, University of Virginia

Devan V. Mehrotra, Merck Research Laboratories

## 2018 ENAR Student Awards

Scarlett L. Bellamy, Drexel University

## ENAR Fostering Diversity in Biostatistics Workshop

Portia D. Exum, SAS Institute Inc.

Felicia R. Simpson, Winston-Salem State University

## ENAR Junior Researchers Workshop

Howard Chang, Emory University

Betsy Ogburn, Johns Hopkins University

Jessica M. Franklin, Brigham and Women's Hospital

David Vock, University of Minnesota

Chris Slaughter, Vanderbilt University

## ENAR Executive Team

Kathy Hoskins, ENAR Executive Director

Katie Earley, ENAR Deputy Director

Laura Stapleton, ENAR Administrative Assistant

Jason Pautler, ENAR Creative Director



# WELCOME TO PHILADELPHIA

Known as “America’s Birthplace,” Philadelphia is a multifaceted city offering visitors fantastic food, historic landmarks and beautiful parks. Ranked number two in 2017 by U.S. News & World Report’s “Best Places to Visit in the USA,” some of Philadelphia’s famous former residents include Benjamin Franklin, Margaret Mead, Grace Kelly and Will Smith. The original city planners designed Philadelphia using a grid system, and today visitors can explore the city easily by foot, bus, subway or trolley. The site of America’s first zoo, public hospital and once the national capital, the residents of this city have been trailblazers since William Penn established Philadelphia in 1682.

## LANDMARKS

### Liberty Bell Center

A 15-minute walk from the Marriott, the Liberty Bell Center is a unique attraction that provides visitors a 360-degree view of the famous Liberty Bell. Once hung from the former Pennsylvania State House, this symbol of freedom has been housed in this interactive center since 2003. Millions of visitors each year visit to snap a photo and learn more about how this iconic bell was adopted by the abolitionist and suffrage movements.

### Philadelphia Zoo

Only three miles from the Marriott, the Philadelphia Zoo is home to more than 1,000 animals. Established in 1859, America’s first zoo welcomes more than 1 million visitors a year and offers activities for all ages. Zoo360 provides guests an opportunity to explore the zoo and the habitats of gorillas, meerkats and tigers like never before. KidZooU provides educational activities for many of the younger guests.

### The “Rocky” Statue/“Rocky” Steps

Located only two miles from the Marriott at the Philadelphia Museum of Art, this local landmark was the setting of one of the most iconic movie scenes of all time. Considered a “rite of passage” for Philadelphia visitors, run up the 72 steps and experience a beautiful view of Philadelphia and on your way down, snap a photo with the “Rocky” statue.

### Eastern State Penitentiary

Once considered one of America’s most famous prisons, Philadelphia’s Eastern State Penitentiary formerly housed such inmates as bank robber Slick Willie Sutton and gangster Al Capone. No longer in operation as a prison, Eastern State is now a museum offering self-guided tours and hands-on activities. A variety of exhibits are included with admission, which discuss such topics as the prison system today and the first prison synagogue.

## HISTORY & ART MUSEUMS

### Independence Hall

A focal point of this historic city, Independence Hall was the site of the signing of the Declaration of Independence and the U.S. Constitution. Built in 1732 as the Pennsylvania State House, this building is now managed by the National Park Service. Tours are offered of the Assembly Room and Courtroom of the Pennsylvania Supreme Court, among others, which are decorated with 18th century furniture to allow visitors to truly immerse themselves in American history.

### The Franklin Institute

A 20-minute walk from the Marriott, The Franklin Institute has become one of the nation’s premier science museums. Established in 1824 and named after Benjamin Franklin, The Franklin Institute is the most visited museum in Pennsylvania, with more than 1 million visitors every year. Visitors of all ages will enjoy the institute’s interactive exhibits, such as the two-story human heart, as part of a day of learning and fun.

### The Betsy Ross House

Just one mile from the Marriott, the Betsy Ross House offers visitors a tour of the historical house and an opportunity to learn more about the woman, who made the first American flag. The cost of admission includes a self-guided audio tour of the house, interaction with colonial reenactors and hands-on activities.

### Philadelphia Museum of Art

Since its founding in 1876, the Philadelphia Museum of Art is one of the largest art museums in the country. Work by such artists as Paul Cézanne, Rogier van der Weyden and Marcel Duchamp are showcased in the museum’s 200+ galleries. The museum offers extended hours on Friday nights, as well as live music, cocktails and small plates.

### Rodin Museum

The Rodin Museum, located on the Benjamin Franklin Parkway and a 15-minute walk from the Philadelphia Museum of Art, houses the most comprehensive collection of sculptor Auguste Rodin’s works outside Paris. The museum was founded in 1929 and is listed on the Philadelphia register of historic places. The most famous of Rodin’s sculptures, The Thinker (1880-1882), stands in the courtyard leading to this magnificent museum.

### The Barnes

The Barnes Foundation, located on the Ben Franklin Parkway, boasts an impressive and eclectic array of well-known impressionist and modernist paintings. Founded by Albert C. Barnes, who made his fortune in the 1920’s from the discovery of an antiseptic used widely in hospitals, the collection was first housed in his private home in the suburbs of Philadelphia but is now in a modern space filled with natural light. This gem is a must see.



## PARKS & NEIGHBORHOODS

### Dilworth Park

Just a 5-minute walk from the Marriott, Dilworth Park is one of the newer parks in Philadelphia, having opened in 2014. Dilworth Park has activities year-round, and in the winter months, a skating rink is open to the public. Explore America's Garden Capital Maze, which showcases more than 30 gardens from the Philadelphia-metro area. A café offering breakfast, lunch, dinner and drinks is onsite as well.

### Fairmont Park

Covering more than 2,000 acres, Fairmont Park is packed with sites and attractions not found in most other parks. From the Carousel House to the Shofuso Japanese House and Garden to the Turtle Rock Lighthouse, Fairmont Park provides visitors with a full day of sightseeing. Don't miss out on taking the trolley tour and explore the 18th and 19th century mansions, which are open to the public.

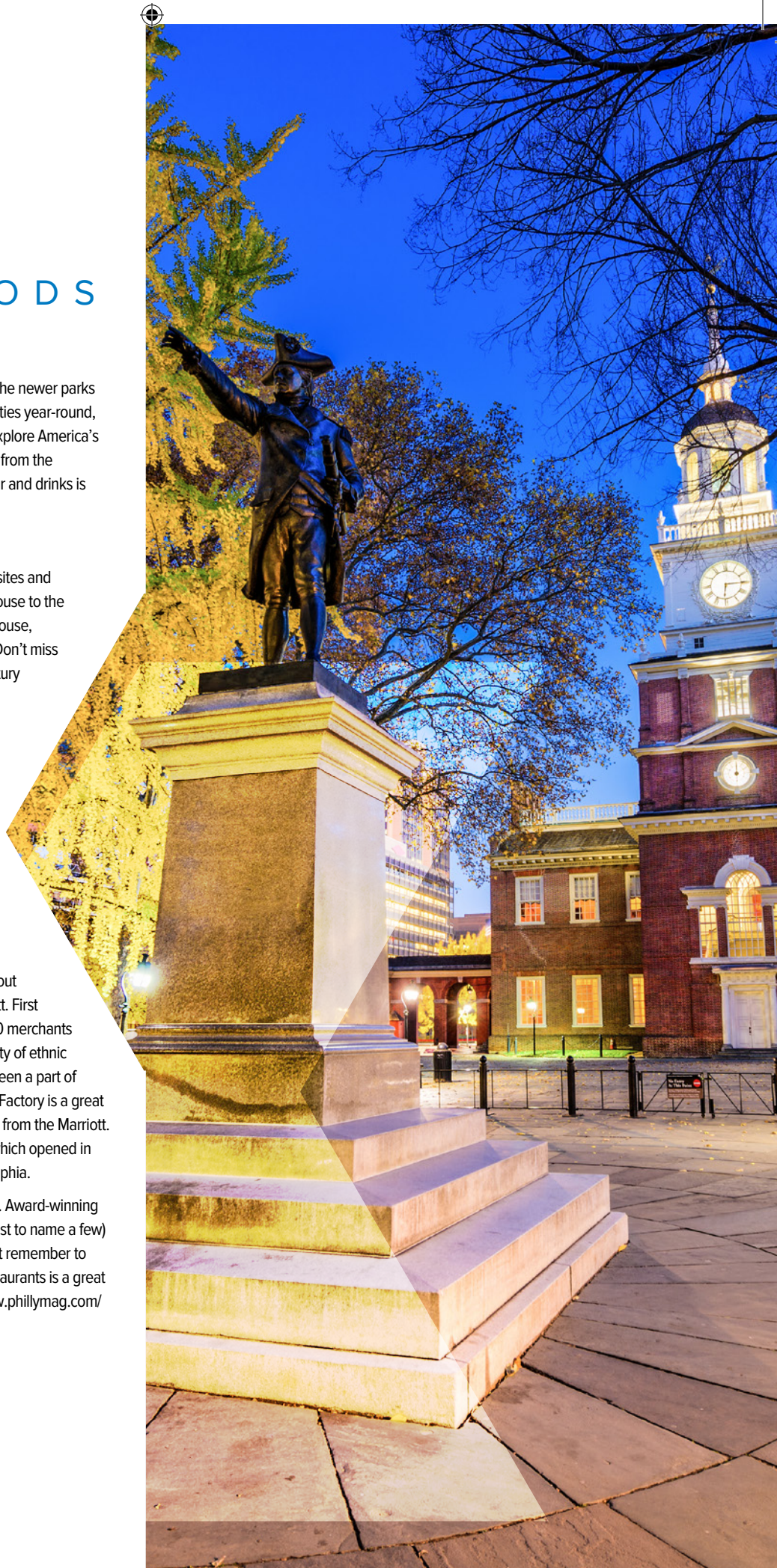
### Center City

Located in the heart of Philadelphia, Center City includes the neighborhoods of Chinatown, Logan Square, the French Quarter, Old City and Society Hill to name a few. Some of the best shopping and historical sites Philadelphia has to offer are located in Center City. Thanks to Philadelphia's vast subway system, visitors can explore Center City with ease.

### Local Cuisine and Restaurants

To get a taste of what the Philly food scene has to offer, check out Reading Terminal Market, just a 2-minute walk from the Marriott. First opened in 1893, today this historic market houses more than 80 merchants offering fresh produce, meats, fish, groceries, as well as a variety of ethnic foods. Not known just for its cheesesteaks, soft pretzels have been a part of the Philadelphia food scene for generations. The Philly Pretzel Factory is a great spot to try a hand-twisted creation, located less than half a mile from the Marriott. Another landmark worth visiting is McGillin's Olde Ale House, which opened in 1860 and is the oldest continuously operating tavern in Philadelphia.

Philadelphia's reputation as a town for "foodies" is well earned. Award-winning restaurants, such as Zahav, Vernick, Vetri, Suraya and Laurel (just to name a few) will make your visit to Philadelphia that much more special. Just remember to make reservations! Philadelphia Magazine's list of 50 Best Restaurants is a great resource for all types of palates, diets and budgets: <https://www.phillymag.com/foobooz/50-best-restaurants/>





# PRESIDENTIAL INVITED SPEAKER



## Francesca Dominici, Ph.D

Clarence James Gamble Professor of Biostatistics, Population and Data Science,  
Harvard T.H. Chan School of Public Health  
Co-Director, Data Science Initiative, Harvard University

### A Particulate Solution: Data Science in the Fight to Stop Air Pollution and Climate Change

What if I told you I had evidence of a serious threat to American national security – a terrorist attack in which a jumbo jet will be hijacked and crashed every 12 days. Thousands will continue to die unless we act now. This is the question before us today – but the threat doesn't come from terrorists. The threat comes from climate change and air pollution.

We have developed an artificial neural network model that uses on-the-ground air-monitoring data and satellite-based measurements to estimate daily pollution levels across the continental U.S., breaking the country up into 1-square-kilometer zones. We have paired that information with health data contained in Medicare claims records from the last 12 years, and for 97% of the population ages 65 or older. We have developed statistical methods and computationally efficient algorithms for the analysis over 460 million health records.

Our research shows that short and long term exposure to air pollution is killing thousands of senior citizens each year. This data science platform is telling us that federal limits on the nation's most widespread air pollutants are not stringent enough.

This type of data is the sign of a new era for the role of data science in public health, and also for the associated methodological challenges. For example, with enormous amounts of data, the threat of unmeasured confounding bias is amplified, and causality is even harder to assess with observational studies. These and other challenges will be discussed.

### Biography

Francesca Dominici received her PhD in Statistics from the University of Padua, Italy, in 1997. From 1999 to 2009 she was a Professor at the Bloomberg School of Public Health at Johns Hopkins University. In 2009 she moved to the Harvard T.H. Chan School of Public Health as a tenured Professor of Biostatistics and was appointed Associate Dean of Information Technology in 2011. In Fall 2013, she was appointed Senior Associate Dean for Research and in February 2017, she was appointed as co-director of the Harvard Data Science Initiative.

Dr. Dominici's research has focused on the development of statistical methods for the analysis of large and complex data. She is a passionate data scientist; her expertise is in the development of statistical methods for the analysis of large, messy data and for combining information across heterogeneous data sources. She leads several interdisciplinary groups of scientists with the ultimate goal of addressing important questions in environmental health science, climate change, comparative effectiveness research in cancer, and health policy.

In her current role as co-director of the Data Science Initiative at Harvard University, Dr. Dominici is building on the collaborations that already exist across the University to foster a rich and cohesive data science community, bringing together scholars from across disciplines and schools. In her role as Senior Associate Dean for Research, she led advancements to optimize the health of the Harvard Chan School's research enterprise, and led the School's Office of Research Strategy and Development flagship faculty grant-writing short course. Dr. Dominici has personally contributed scientific leadership in the submission of myriad pioneering proposals designed to advance scientific innovation and the field of public health at large through data science. In addition to her research interests and administrative leadership roles, she has demonstrated a career-long commitment to promoting diversity in academia.

Dr. Dominici is an elected fellow of the American Statistical Association. She has received a number of prestigious awards and honors, including the 2016 Janet L. Norwood Award, 2015 Florence Nightingale David Award, 2009 Diversity Recognition Award from Johns Hopkins University, 2007 Gertrude Cox Award, and even a 1998 ENAR student award!



# PROGRAM SUMMARY

SATURDAY, MARCH 23	
8:30 a.m.—8:00 p.m. <i>Franklin Hall 2, 4th Floor</i>	Workshop for Junior Researchers
3:30 p.m.—5:30 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Conference Registration
SUNDAY, MARCH 24	
7:30 a.m.—6:30 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Conference Registration
8:00 a.m.—12:00 p.m.	Short Courses
SC4 <i>Franklin Hall 4, 4th Floor</i>	StatTag for Connecting R, SAS, and Stata to Word: A Practical Approach to Reproducibility
SC5 <i>Franklin Hall 13, 4th Floor</i>	Personalized Medicine: Subgroup Identification in Clinical Trials
8:00 a.m.—5:00 p.m.	Short Courses
SC1 <i>Franklin Hall 1, 4th Floor</i>	Bayesian Inference and Clinical Trial Designs Using Historical Data
SC2 <i>Franklin Hall 2, 4th Floor</i>	Big Data, Data Science and Deep Learning for Statistician
SC3 <i>Franklin Hall 3, 4th Floor</i>	Analysis of Medical Cost Data: Statistical and Econometric Tools
12:30 p.m. – 5:30 p.m. <i>Salon A, 5th Floor</i>	Fostering Diversity in Biostatistics Workshop
1:00 p.m.—5:00 p.m.	Short Courses
SC6 <i>Franklin Hall 13, 4th Floor</i>	Design of Matched Studies with Improved Internal and External Validity
SC7 <i>Franklin Hall 4, 4th Floor</i>	Smart Simulations with SAS and R
3:00 p.m.—6:00 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Exhibits Open
4:30 p.m.—7:00 p.m. <i>Conference Suite III, 3rd Floor</i>	ENAR Executive Committee Meeting
4:00 p.m.—6:30 p.m. <i>Room 402-403, 4th Floor</i>	Career Placement Service
7:30 p.m.—8:00 p.m. <i>Salons E-F, 5th Floor</i>	New Member Reception
8:00 p.m.—11:00 p.m. <i>Salons E-F, 5th Floor</i>	Opening Mixer and Poster Session
1.	Posters: Variable Subset Selection

2.	Posters: Survival Analysis/Competing Risks
3.	Posters: Machine Learning
4.	Posters: Personalized Medicine
5.	Posters: Cancer Applications
6.	Posters: Clinical Trials
7.	Posters: Diagnostics/Agreement
8.	Posters: Adaptive Design/Experimental Design
9.	Posters: Bayesian Methods
10.	Posters: Bayesian Omics/Latent Bayes Models
11.	Posters: Genomics/Proteomics
12.	Posters: Functional Data/High Dimensional
13.	Posters: Spatial/Temporal Modeling
14.	Posters: Public Health/Surveys/EHR
MONDAY, MARCH 25	
7:30 a.m.—5:00 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Conference Registration
7:30 a.m. – 5:00 p.m. <i>Conference Room III, 3rd Floor</i>	Speaker Ready Room
8:30 a.m.—5:30 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Exhibits Open
8:30 a.m.—10:15 a.m. <i>Salon A, 5th Floor</i>	TUTORIAL T1: An Introduction to Causal Effect Estimation with Examples Using SAS Software
	Scientific Program
15. <i>Franklin Hall 13, 4th Floor</i>	Novel Neuroimaging Methods from Processing to Analysis
16. <i>Salon B, 5th Floor</i>	Statistical Challenges and Opportunities for Analysis of Large-Scale Omics Data
17. <i>Salon C, 5th Floor</i>	Longitudinal and Functional Models for Predicting Clinical Outcomes

# PROGRAM SUMMARY (CONTINUED)

18. <i>Salon D, 5th Floor</i>	Recent Bayesian Methods for Causal Inference
19. <i>Franklin Hall 3, 4th Floor</i>	New Methods for Cost-Effectiveness Analysis in Health Policy Research
20. <i>Franklin Hall 2, 4th Floor</i>	Foundations of Statistical Inference in the Era of Machine Learning
21. <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Clinical Trials: Cancer Applications and Survival Analysis
22. <i>405, 4th Floor</i>	Contributed Papers: Multiple Testing
23. <i>406, 4th Floor</i>	Contributed Papers: Clustered Data Methods
24. <i>302-303, 3rd Floor</i>	Contributed Papers: Genome Wide Association Studies and Other Genetic Studies
25. <i>407, 4th Floor</i>	Contributed Papers: Time Series
9:30 a.m. – 4:30 p.m. <i>Room 402-403, 4th Floor</i>	Career Placement Service
10:15 a.m.—10:30 a.m. <i>Franklin Hall Foyer, 4th Floor</i>	Refreshment Break with Our Exhibitors
10:30 a.m.—12:15 p.m. <i>Salon A, 5th Floor</i>	<b>TUTORIAL</b> T2: Building Effective Data Visualizations with ggplot2
	<b>Scientific Program</b>
26. <i>Franklin Hall 2, 4th Floor</i>	Douglas Altman: A Consummate Medical Statistician
27. <i>Salon B, 5th Floor</i>	Statistical Advances for Emerging Issues in Human Microbiome Researches
28. <i>Franklin Hall 13, 4th Floor</i>	Wearable Technology in Large Observational Studies
29. <i>Salon C, 5th Floor</i>	Causal Inference with Non-Ignorable Missing Data: New Developments in Identification and Estimation
30. <i>Salon D, 5th Floor</i>	Advanced Development in Joint Modeling and Risk Prediction
31. <i>Franklin Hall 3, 4th Floor</i>	Classification and Variable Selection under Asymmetric Loss
32. <i>Franklin Hall 1, 4th Floor</i>	Speed Posters: High-Dimensional Data/Omics
33. <i>405, 4th Floor</i>	Contributed Papers: Personalized Medicine
34. <i>302-303, 3rd Floor</i>	Contributed Papers: Epidemiologic Methods

35. <i>406, 4th Floor</i>	Contributed Papers: Statistical Genetics: Single-Cell Sequencing/Transcriptomic Data
36. <i>407, 4th Floor</i>	Contributed Papers: Machine Learning and Testing with High Dimensional Data
37. <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Spatio-Temporal Modeling
12:15 p.m.—1:30 p.m. <i>Salon F, 5th Floor</i>	Roundtable Luncheons
12:30 p.m.—4:30 p.m. <i>404, 4th Floor</i>	Regional Advisory Board (RAB) Luncheon Meeting (by Invitation Only)
1:45 p.m.—3:30 p.m. <i>Salon A, 5th Floor</i>	<b>TUTORIAL</b> T3: Meta-Analysis of Clinical Trials: Effects-at-Random or Studies-at-Random?
	<b>Scientific Program</b>
38. <i>Salon B, 5th Floor</i>	Emerging Statistical Issues and Methods for Integrating Multi-Domain mHealth Data
39. <i>Franklin Hall 13, 4th Floor</i>	Causal Inference with Difference-in-Differences and Regression Discontinuity Designs
40. <i>Franklin Hall 2, 4th Floor</i>	Statistical Challenges in Synthesizing Electronic Healthcare Data
41. <i>Franklin Hall 3, 4th Floor</i>	Using Historical Data to Inform Decisions in Clinical Trials: Evidence based Approach in Drug Development
42. <i>Salon C, 5th Floor</i>	Statistical Innovations in Single-Cell Genomics
43. <i>Salon D, 5th Floor</i>	Advances in Statistical Methods for Surveillance Data of Infectious Diseases
44. <i>Franklin Hall 1, 4th Floor</i>	Speed Posters: Spatio-Temporal Modeling/ Longitudinal Data/ Survival Analysis
45. <i>405, 4th Floor</i>	Contributed Papers: Biomarkers
46. <i>406, 4th Floor</i>	Contributed Papers: Bayesian Modeling and Variable Selection
47. <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Functional Data Applications and Methods
48. <i>302-303, 3rd Floor</i>	Contributed Papers: Machine Learning and Statistical Relational Learning
49. <i>407, 4th Floor</i>	Contributed Papers: Interval-Censored and Multivariate Survival Data
3:30 p.m.—3:45 p.m. <i>Franklin Hall Foyer, 4th Floor</i>	Refreshment Break with Our Exhibitors



## PROGRAM SUMMARY (CONTINUED)

<b>3:45 p.m.—5:30 p.m.</b> <i>Salon A, 5th Floor</i>	<b>TUTORIAL</b> <b>T4: Modern Multiple Imputation</b>
	<b>Scientific Program</b>
<b>50.</b> <i>Franklin Hall 2, 4th Floor</i>	Understanding the Complexity and Integrity of Clinical Trial Data
<b>51.</b> <i>Franklin Hall 13, 4th Floor</i>	Replicability in Big Data Precision Medicine
<b>52.</b> <i>Salon B, 5th Floor</i>	Computationally-Intensive Bayesian Techniques for Biomedical Data: Recent Advances
<b>53.</b> <i>Salon C, 5th Floor</i>	Multivariate Functional Data Analysis with Medical Applications
<b>54.</b> <i>Salon D, 5th Floor</i>	Methods for Examining Health Effects of Exposure to the World Trade Center Attack and Building Collapse
<b>55.</b> <i>Franklin Hall 3, 4th Floor</i>	Regression, Mediation, and Graphical Modeling Techniques for Microbiome Data
<b>56.</b> <i>Franklin Hall 1, 4th Floor</i>	Speed Posters: EHR Data, Epidemiology, Personalized Medicine, Clinical Trials
<b>57.</b> <i>405, 4th Floor</i>	Contributed Papers: Agreement Measures and Diagnostics
<b>58.</b> <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Variable Selection
<b>59.</b> <i>302-303, 3rd Floor</i>	Contributed Papers: Causal Inference
<b>60.</b> <i>406, 4th Floor</i>	Contributed Papers: Genetic Effects/Heritability
<b>61.</b> <i>407, 4th Floor</i>	Contributed Papers: Computational Methods and Massive Data Sets
<b>5:30 p.m. – 6:30 p.m.</b> <i>Salon F, 5th Floor</i>	<b>CENS Networking Mixer</b>
<b>6:30 p.m.—7:30 p.m.</b> <i>Salon D, 5th Floor</i>	<b>President's Reception (by Invitation Only)</b>
<b>TUESDAY, MARCH 26</b>	
<b>7:30 a.m.—5:00 p.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	<b>Conference Registration</b>
<b>7:30 a.m. – 5:00 p.m.</b> <i>Conference III, 3rd Floor</i>	<b>Speaker Ready Room</b>
<b>8:30 a.m.—5:30 p.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	<b>Exhibits Open</b>
<b>9:30 a.m.—3:30 p.m.</b> <i>401-402, 4th Floor</i>	<b>Career Placement Service</b>

<b>8:30 a.m. – 10:15 a.m.</b>	<b>Scientific Program</b>
<b>62.</b> <i>Franklin Hall 1, 4th Floor</i>	Recent Advances in Bayesian Network Meta-Analysis
<b>63.</b> <i>Franklin Hall 2, 4th Floor</i>	Statistical Methods to Support Valid and Efficient use of Electronic Health Records Data
<b>64.</b> <i>Franklin Hall 13, 4th Floor</i>	Recent Advances in the Analysis of Time-to-Event Outcomes Subject to a Terminal Event
<b>65.</b> <i>Salon B, 5th Floor</i>	Recent Advances in Statistical Methods for Precision Medicine
<b>66.</b> <i>Salon C, 5th Floor</i>	Challenges and Advances in Wearable Technology
<b>67.</b> <i>Salon D, 5th Floor</i>	Statistical Modeling in Cell Biology
<b>68.</b> <i>405, 4th Floor</i>	Contributed Papers: Diagnostics, ROC, and Risk Prediction
<b>69.</b> <i>406, 4th Floor</i>	Contributed Papers: Microbiome Data: Finding Associations and Testing
<b>70.</b> <i>Franklin Hall 3, 4th Floor</i>	Contributed Papers: Comparative Effectiveness, Clustered and Categorical Data
<b>71.</b> <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Causal Inference and Measurement Error
<b>72.</b> <i>407, 4th Floor</i>	Contributed Papers: Genomics, Proteomics, or Other Omics
<b>73.</b> <i>302-303, 3rd Floor</i>	Contributed Papers: Imaging Methods
<b>10:15 a.m.—10:30 a.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	<b>Refreshment Break with Our Exhibitors</b>
<b>10:30 a.m.—12:15 p.m.</b>	<b>Scientific Program</b>
<b>74.</b> <i>Salons E-F, 5th Floor</i>	Presidential Invited Address
<b>12:30 p.m.—4:30 p.m.</b> <i>404, 4th Floor</i>	<b>Regional Committee Luncheon Meeting (by Invitation Only)</b>
<b>1:45 p.m.—3:30 p.m.</b> <i>Salon A, 5th Floor</i>	<b>TUTORIAL</b> <b>T5: A Primer on Python for Statistical Programming and Data Science</b>
	<b>Scientific Program</b>
<b>75.</b> <i>Franklin Hall 1, 4th Floor</i>	Resource Efficient Study Designs for Observational and Correlated Data
<b>76.</b> <i>Franklin Hall 2, 4th Floor</i>	Recent Advances in the Study of Interaction

## PROGRAM SUMMARY (CONTINUED)

<b>77.</b> <i>Salon B, 5th Floor</i>	Expanding Rank Tests: Estimates, Confidence Intervals, Modeling, and Applications
<b>78.</b> <i>Franklin Hall 13, 4th Floor</i>	Novel Statistical Methods to Analyze Self-Reported Outcomes Subject to Recall Error in Observational Studies
<b>79.</b> <i>Salon C, 5th Floor</i>	Statistical Advance in Human Microbiome Data Analysis
<b>80.</b> <i>Salon D, 5th Floor</i>	Statistical Mediation Analysis for High-Dimensional Data
<b>81.</b> <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Prediction and Prognostic Modeling
<b>82.</b> <i>405, 4th Floor</i>	Contributed Papers: Adaptive Designs for Clinical Trials
<b>83.</b> <i>406, 4th Floor</i>	Contributed Papers: Bayesian Approaches to Surveys and Spatio-Temporal Modeling
<b>84.</b> <i>407, 4th Floor</i>	Contributed Papers: Causal Effects with Propensity Scores/Weighting/Matching
<b>85.</b> <i>302-303, 3rd Floor</i>	Contributed Papers: Meta-Analysis
<b>86.</b> <i>Franklin Hall 3, 4th Floor</i>	Contributed Papers: Imaging Applications and Testing
<b>3:30 p.m.—3:45 p.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	<b>Refreshment Break with Our Exhibitors</b>
<b>3:45 p.m.—5:30 p.m.</b> <i>Salon A, 5th Floor</i>	<b>TUTORIAL</b> <b>T6: Analysis of Patient-Reported Outcomes</b>
	<b>Scientific Program</b>
<b>87.</b> <i>Franklin Hall 1, 4th Floor</i>	Methodological Challenges and Opportunities in Mental Health Research
<b>88.</b> <i>Salon B, 5th Floor</i>	Novel Approaches for Group Testing for Estimation in Biostatistics
<b>89.</b> <i>Franklin Hall 13, 4th Floor</i>	Adaptive and Bayesian Adaptive Design in Bioequivalence and Biosimilar Studies
<b>90.</b> <i>Salon C, 5th Floor</i>	Methods to Robustly Incorporate External Data into Genetic Tests
<b>91.</b> <i>Franklin Hall 2, 4th Floor</i>	Developing Collaborative Skills for Successful Careers in Biostatistics and Data Science
<b>92.</b> <i>Salon D, 5th Floor</i>	New Approaches to Causal Inference under Interference: Bringing Methodological Innovations into Practice
<b>93.</b> <i>Franklin Hall 3, 4th Floor</i>	Contributed Papers: Design and Analysis of Clinical Trials

<b>94.</b> <i>405, 4th Floor</i>	Contributed Papers: Semiparametric, Nonparametric, and Empirical Likelihood Models
<b>95.</b> <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Bayesian Approaches to High Dimensional Data
<b>96.</b> <i>406, 4th Floor</i>	Contributed Papers: Functional Data Analysis Methods
<b>97.</b> <i>302-303, 3rd Floor</i>	Contributed Papers: Next Generation Sequencing
<b>98.</b> <i>407, 4th Floor</i>	Contributed Papers: Competing Risks and Cure Models
<b>5:45 p.m.—7:00 p.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	ENAR Business Meeting and Sponsor/Exhibitor Mixer
<b>WEDNESDAY, MARCH 27</b>	
<b>7:30 a.m. – 12:00 noon</b> <i>Conference Suite III, 3rd Floor</i>	<b>Speaker Ready Room</b>
<b>7:30 a.m.—9:00 a.m.</b> <i>411, 4th Floor</i>	<b>Planning Committee (by Invitation Only)</b>
<b>8:00 a.m.—12:30 p.m.</b> <i>Franklin Foyer, 4th Floor</i>	<b>Conference Registration</b>
<b>8:00 a.m.—12:00 p.m.</b> <i>Franklin Hall Foyer, 4th Floor</i>	<b>Exhibits Open</b>
<b>8:30 a.m.—10:15 a.m.</b>	<b>Scientific Program</b>
<b>99.</b> <i>Salon B, 5th Floor</i>	Monitoring Health Behaviors with Multi-Sensor Mobile Technology
<b>100.</b> <i>Salon C, 5th Floor</i>	Current Methods to Address Data Errors in Electronic Health Records
<b>101.</b> <i>Franklin Hall 1, 4th Floor</i>	Finding the Right Academic Fit: Experiences from Faculty across the Academic Spectrum
<b>102.</b> <i>Franklin Hall 2, 4th Floor</i>	Novel Integrative Omics Approaches for Understanding Complex Human Diseases
<b>103.</b> <i>Salon D, 5th Floor</i>	Teaching Data Science through Case-Studies
<b>104.</b> <i>Franklin Hall 13, 4th Floor</i>	Nonconvex Optimization and Biological Applications
<b>105.</b> <i>302-303, 3rd Floor</i>	Contributed Papers: Biopharmaceutical Research and Clinical Trials
<b>106.</b> <i>405, 4th Floor</i>	Contributed Papers: Missing Data
<b>107.</b> <i>Franklin Hall 4, 4th Floor</i>	Contributed Papers: Bayesian Computational and Modeling Methods



## PROGRAM SUMMARY (CONTINUED)

<b>108.</b> 406, 4th Floor	Contributed Papers: Causal Effect Modeling (Mediation/Variable Selection/Longitudinal)
<b>109.</b> Franklin Hall 3, 4th Floor	Contributed Papers: Microbiome Data Analysis with Zero Inflation and/or Model Selection
<b>110.</b> 407, 4th Floor	Contributed Papers: Recurrent Events or Multiple Time-to-Event Data
<b>10:15 a.m.—10:30 a.m.</b> Franklin Hall Foyer, 4th Floor	<b>Refreshment Break with Our Exhibitors</b>
<b>10:30 a.m.—12:15 p.m.</b>	<b>Scientific Program</b>
<b>111.</b> Franklin Hall 1, 4th Floor	Individualized Evidence for Medical Decision Making: Principles and Practices
<b>112.</b> Franklin Hall 2, 4th Floor	Some New Perspectives and Developments for Data Integration in the Era of Data Science
<b>113.</b> Salon B, 5th Floor	Bayesian Methods for Spatial and Spatio-Temporal Modeling of Health Data
<b>114.</b> Salon C, 5th Floor	Recent Advances in Causal Inference for Survival Analysis
<b>115.</b> Salon D, 5th Floor	Novel Statistical Methods for Analysis of Microbiome Data
<b>116.</b> Franklin Hall 13, 4th Floor	New Developments in Nonparametric Methods for Covariate Selection
<b>117.</b> Franklin Hall 3, 4th Floor	Contributed Papers: Dynamic Treatment Regimens and Experimental Design
<b>118.</b> 405, 4th Floor	Contributed Papers: Hypothesis Testing and Sample Size Calculation
<b>119.</b> 302-303, 3rd Floor	Contributed Papers: Measurement Error
<b>120.</b> 406, 4th Floor	Contributed Papers: Environmental and Ecological Applications
<b>121.</b> Franklin Hall 4, 4th Floor	Contributed Papers: Statistical Methods for High Dimensional Data
<b>122.</b> 407, 4th Floor	Contributed Papers: Longitudinal Data and Joint Models of Longitudinal and Survival Data



# SCIENTIFIC PROGRAM

SUNDAY, MARCH 24

SUNDAY, MARCH 24

8:00—11:00 P.M.

## POSTER PRESENTATIONS

Salons E-F, 5th Floor

### 1. POSTERS: VARIABLE SUBSET SELECTION

**SPONSOR:** ENAR

#### 1a. Total Variation Denoising of Coefficient Functionals in the Additive Complementary Log-Log Survival Model

Hao Sun\* and Brent A. Johnson, University of Rochester

#### 1b. Bayesian Variable Selection for High-Dimensional Data with Ordinal Responses

Yiran Zhang\* and Kellie J. Archer, The Ohio State University

#### 1c. Selecting Appropriate Probabilistic Models for Microbiome Data Analysis

Hani Aldirawi\* and Jie Yang, University of Illinois at Chicago; Ahmed A. Metwally, Stanford University

#### 1d. On the Must-Be of Variable Selection in Biomedical Research

Bokai Wang\* and Changyong Feng, University of Rochester

#### 1e. An Estimation of Average Treatment Effect using Adaptive Lasso and Doubly Robust Estimator

Wataru Hongo\*, Shuji Ando, Jun Tsuchida and Takashi Sozu, Tokyo University of Science

### 2. POSTERS: SURVIVAL ANALYSIS/COMPETING RISKS

**SPONSOR:** ENAR

#### 2a. Augmented Double Inverse-Weighted Estimation of Difference in Restricted Mean Lifetimes using Observational Data Subject to Dependent Censoring

Qixing Liang\* and Min Zhang, University of Michigan

#### 2b. Single-Index Models with Transformation Models for Optimal Treatment Regimes

Jin Wang\* and Danyu Lin, University of North Carolina, Chapel Hill

#### 2c. Univariate Gradient Statistic for Marginal Cure Rate Model with High-Dimensional Covariates

Jennifer L. Delzeit\*, Jianfeng Chen and Wei-Wen Hsu, Kansas State University; David Todem, Michigan State University; KyungMann Kim, University of Wisconsin, Madison

#### 2d. Integrative Survival Analysis with Uncertain Event Times in Application to a Suicide Risk Study

Wenjie Wang\*, Robert Aseltine, Kun Chen and Jun Yan, University of Connecticut

#### 2e. Nonparametric Estimation of the Joint Distribution of a Survival Time and Mark Variable in the Presence of Dependent Censoring

Busola Sanusi\*, Jianwen Cai and Michael G. Hudgens, University of North Carolina, Chapel Hill

#### 2f. Group Variable Screening for Clustered Multivariate Survival Data

Natasha A. Sahr\*, St. Jude Children's Research Hospital; Kwang Woo Ahn and Soyoung Kim, Medical College of Wisconsin

#### 2g. Semiparametric Regression on Cumulative Incidence Function with Interval-Censored Competing Risks Data and Missing Event Type

Jun Park\*, Giorgos Bakoyannis, Ying Zhang and Constantin T. Yiannoutsos, Indiana University

#### 2h. Latent Class Regression Modeling of Competing Risks Data

Teng Fei\*, Emory Rollins School of Public Health; John Hanfelt, Emory Rollins School of Public Health and Emory Alzheimer's Disease Research Center; Limin Peng, Emory Rollins School of Public Health

#### 2i. The Use of Repeated Measurements for Dynamic Cardiovascular Disease Prediction: The Application of Joint Model in the Lifetime Risk Pooling Project

Yu Deng\*, Yizhen Zhong and Abel Kho, Northwestern University; Lei Liu, Washington University School of Medicine; Norrina Allen, John Wilkins, Kiang Liu, Donald Lloyd-Jones and Lihui Zhao, Northwestern University

### 3. POSTERS: MACHINE LEARNING

**SPONSOR:** ENAR

#### 3a. Learning Image with Gaussian Process Regression and Application to Classification

Tahmidul Islam\*, University of South Carolina

#### 3b. The Models Underlying Word2Vec, a Natural Language Processing Algorithm, and their Relationship to Traditional Statistical Multivariate Methods

Brian L. Egleston\*, Fox Chase Cancer Center, Temple University Health System; Tian Bai and Slobodan Vucetic, Temple University

#### 3c. A Two-Step Clustering Algorithm for Clustering Data with Mixed Variable Types

Shu Wang\*, Jonathan G. Yabes and Chung-Chou H. Chang, University of Pittsburgh

#### 3d. PseudoNet: Reconstructing Pseudo-Time in Single-Cell RNA-seq Data Using Neural Networks

Justin Lakkis\*, University of Pennsylvania; Chenyi Xue, Huize Pan, Sarah B. Trignano and Hanrui Zhang, Columbia University; Nancy Zhang, University of Pennsylvania; Muredach Reilly, Columbia University; Gang Hu, Nankai University; Mingyao Li, University of Pennsylvania

# SCIENTIFIC PROGRAM

SUNDAY, MARCH 24

**3e. Patterns of Comorbidity Preceding Dementia Diagnosis: Findings from the Atherosclerosis Risk in Communities (ARIC) Study Cohort**  
Arkopal Choudhury\*, Anna M. Kucharska-Newton and Michael R. Kosorok, University of North Carolina, Chapel Hill

## 4. POSTERS: PERSONALIZED MEDICINE

**SPONSOR:** ENAR

**4a. A Utility Approach to Individualized Optimal Dose Selection Using Biomarkers**  
Pin Li\*, Matthew Schipper and Jeremy Taylor, University of Michigan

**4b. A Comparison and Assessment of Recently Developed Tree-Based Methods for Subgroup Identification**  
Xinjun Wang\* and Ying Ding, University of Pittsburgh

**4c. A Simultaneous Inference Procedure to Identify Subgroups in Targeted Therapy Development with Time-to-Event Outcomes**  
Yue Wei\* and Ying Ding, University of Pittsburgh

**4d. A Basket Trial Design using Bayesian Model Averaging**  
Akihiro Hirakawa\* and Ryo Sadachi, The University of Tokyo

**4e. Outcome Weighted  $\psi$ -learning for Individualized Treatment Rules**  
Mingyang Liu\*, Xiaotong Shen and Wei Pan, University of Minnesota

**4f. Domain Adaptation Machine Learning for Optimizing Treatment Strategies in Randomized Trials by Leveraging Electronic Health Records**  
Peng Wu\* and Yuanjia Wang, Columbia University

## 5. POSTERS: CANCER APPLICATIONS

**SPONSOR:** ENAR

**5a. Barcoding of Hematopoietic Stem Cells: Application of the Species Problem**  
Siyi Chen\* and Marek Kimmel, Rice University; Katherine Yudeh King, Baylor College of Medicine

**5b. Inferring Clonal Evolution of Tumors from RNA-seq Data**  
Tingting Zhai\*, Jinpeng Liu, Arnold J. Stromberg and Chi Wang, University of Kentucky

**5c. Modeling the Effect of Treatment Timing on Survival with Application to Cancer Screening**  
Wenjia Wang\* and Alexander Tsodikov, University of Michigan

**5d. Molecular Signature Predictive of Survival in Metastatic Cutaneous Melanoma**  
Yuna Kim\* and Issa Zakeri, Drexel University; Sina Nassiri, The Swiss Institute of Bioinformatics (SIB)

**5e. Can a Tumor's Transcriptome Predict Response to Immunotherapy?**  
Shanika A. De Silva\*, Drexel University; Sina Nassiri, The Swiss Institute of Bioinformatics (SIB); Issa Zakeri, Drexel University

**5f. Similarity and Difference between Telomerase Activation and ALT based on the Theory of G-Networks and Stochastic Automata Networks**  
Katie Kyunghyun Lee\* and Marek Kimmel, Rice University

**5g. Optimization of Moment-based Intensity-Modulated Radiation Therapy (IMRT) Treatment Plan**  
Wanxin Chen\* and Abraham Abebe, Temple University

## 6. POSTERS: CLINICAL TRIALS

**SPONSOR:** ENAR

**6a. Two-Stage Adaptive Enrichment Design for Testing an Active Factor**  
Rachel S. Zahigian\*, University of Florida; Aidong Adam Ding, Northeastern University; Samuel S. Wu and Natalie E. Dean, University of Florida

**6b. Covariate Adjustment for Considering Between-Trial Heterogeneity in Clinical Trials using Historical Data for Evaluating the Treatment Efficacy**  
Tomohiro Ohigashi\*, Takashi Sozu and Jun Tsuchida, Tokyo University of Science

**6c. A Comparison of Statistical Methods for Treatment Effect Testing and Estimation with Potential Non-Proportional Hazards**  
Jing Li\*, Indiana University; Qing Li and Amarjot Kaur, Merck & Co., Inc.

**6d. Dose-Finding Method for Molecularly Targeted Agents Incorporating the Relative Dose Intensity in Phase I Oncology Clinical Trials**  
Yuichi Tanaka\*, Takashi Sozu and Akihiro Hirakawa, The University of Tokyo

**6e. Use of Quadratic Inference Function for Estimation of Marginal Intervention Effects in Randomized Trials**  
**WITHDRAWN**  
Hengshi Yu\*, University of Michigan; Fan Li and Elizabeth Louise Turner, Duke University

**6f. Ongoing Clinical Trial Forecast and Monitoring Tool: A Statistical Framework**  
Shijia Bian\*, Biogen; Jignesh Parikh, Sema4; Tanya Cashorali, Biogen and TCB Analytics; Mike Fitzpatrick and Murray Abramson, Biogen; Feng Gao, Biogen

**6g. Parameter Estimation using Influential Exponential Tilting in Case of Data Missing at Random and Non-Ignorable Missing Data**  
Kavita A. Gohil\*, Hani M. Samawi, Haresh Rochani and Lili Yu, Georgia Southern University



# SCIENTIFIC PROGRAM

SUNDAY, MARCH 24

## 7. POSTERS: DIAGNOSTICS/AGREEMENT

**SPONSOR:** ENAR

### 7a. A Non-Inferiority Test for Comparing Two Predictive Values of Diagnostic Tests

Kanae Takahashi\*, Osaka City University; Kouji Yamamoto, Yokohama City University

### 7b. Improving Inference on Discrete Diagnostic Tests Without a Gold Standard

Xianling Wang\* and Gong Tang, University of Pittsburgh

### 7c. Applications of Generalized Kullback-Leibler Divergence as a Measure of Medical Diagnostic and Cut-Point Criterion for K-Stages Diseases

Chen Mo\*, Hani M. Samawi, Jingjing Yin, Haresh D. Rochani, Xinyan Zhang and Jing Kersey, Georgia Southern University

### 7d. A Nonparametric Procedure for Comparing Dependent Kappa Statistics

Hanna Lindner\*, Phyllis Gimotty and Warren Bilker, University of Pennsylvania

### 7e. Evaluating Different Approaches to Classify Patients of Vector Transmitted Viral Infections Using Symptom Information

Ana Maria Ortega-Villa\* and Sally Hunsberger, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Wenjuan Gu and Keith Lumbard, Frederick National Laboratory for Cancer Research sponsored by the National Cancer Institute, National Institutes of Health; Jesús Sepúlveda-Delgado, Hospital Regional de Alta Especialidad Ciudad Salud. Tapachula, Chiapas; Pablo F. Belaunzaran-Zamudio, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico

## 8. POSTERS: ADAPTIVE DESIGN/EXPERIMENTAL DESIGN

**SPONSOR:** ENAR

### 8a. A Utility-based Seamless Phase I/II Trial Design to Identify the Optimal Biological Dose for Targeted and Immune Therapies

Yanhong Zhou\*, J. Jack Lee and Ying Yuan, University of Texas MD Anderson Cancer Center

### 8b. An Adaptive Biomarker-Driven Phase II Design

Jun Yin\*, Mayo Clinic; Daniel Kang, University of Iowa; Qian Shi, Mayo Clinic

### 8c. A Signature Enrichment Design with Bayesian Adaptive Randomization for Cancer Clinical Trials

Fang Xia\*, University of Texas MD Anderson Cancer Center; Stephen L. George, Duke University School of Medicine; Jing Ning, Liang Li and Xuelin Huang, University of Texas MD Anderson Cancer Center

### 8d. Blinded Sample Size Re-Estimation in Comparative Clinical Trials with Over-Dispersed Count Data: Incorporation of Misspecification of the Variance Function

Masataka Igeta\*, Hyogo College of Medicine; Shigeyuki Matsui, Nagoya University School of Medicine

### 8e. An Exploration of Optimal Design Robustness in Nonlinear Models

Ryan Jarrett\* and Matthew S. Shotwell, Vanderbilt University

## 9. POSTERS: BAYESIAN METHODS

**SPONSOR:** ENAR

### 9a. Constructing a Prior for the Correlation Coefficient Using Expert Elicitation

Divya R. Lakshminarayanan\* and John W. Seaman Jr., Baylor University

### 9b. Consistent Bayesian Joint Variable and DAG Selection in High Dimensions

Xuan Cao\*, University of Cincinnati; Kshitij Khare and Malay Ghosh, University of Florida

### 9c. Bayesian Approach for Joint Modelling of Longitudinal and Time to Event Data

Zeynep Atli\*, Mimar Sinan Fine Arts University; Mithat Gönen, Memorial Sloan Kettering Cancer Center, Gülay Basarir, Mimar Sinan Fine Arts University

### 9d. Rational Determination of the Borrowing Rate in Bayesian Power Prior Models

Mario Nagase\*, AstraZeneca; Shinya Ueda, Mitsuo Higashimori and Katsuomi Ichikawa, AstraZenecaKK; Jim Dunyak and Nidal Al-Huniti, AstraZeneca

### 9e. A Semiparametric Approach for Estimating a Bacterium's Wild-Type Distribution: Accounting for Contamination and Measurement Error (BayesACME)

Will A. Eagan\* and Bruce A. Craig, Purdue University

### 9f. A Bayesian Mixture Model to Estimate the Effect of an Ordinal Predictor

Emily Roberts\* and Lili Zhao, University of Michigan

## 10. POSTERS: BAYESIAN OMICS/LATENT BAYES MODELS

**SPONSOR:** ENAR

### 10a. Inferring Gene Networks with Global-Local Shrinkage Rules

Viral V. Panchal\* and Daniel F. Linder, Augusta University

### 10b. Bayesian Kinetic Modeling for Tracer-Based Metabolomic Data

Xu Zhang\*, Andrew N. Lane, Arnold Stromberg, Teresa W-M. Fan and Chi Wang, University of Kentucky

# SCIENTIFIC PROGRAM

SUNDAY, MARCH 24

## 10c. A Bayesian Approach for Flexible Clustering of Microbiome Data

Yushu Shi\*, Liangliang Zhang, Kim-Anh Do, Robert Jenq and Christine Peterson, University of Texas MD Anderson Cancer Center

## 10d. Latent Scale Prediction Model for Network Valued Covariates

Xin Ma\*, Suprateek Kundu and Jennifer Stevens, Emory University

## 10e. Bayesian Estimation in Latent Variable Analysis in Mplus and WinBUGS

Jinxiang Hu, Lauren Clark\* and Byron Gajewski, University of Kansas Medical Center

## 10f. A Bayesian Factor Model for Healthcare Rankings: Applications in Estimating Composite Measures of Quality

Stephen Salerno\*, Lili Zhao and Yi Li, University of Michigan

## 10g. A Bayesian Latent Variable Model for Inflammatory Markers and Birth Outcomes in Seychelles

Alexis E. Zavez\*, University of Rochester; Alison J. Yeates, Ulster University; Edwin van Wijngaarden and Sally W. Thurston, University of Rochester

## 11. POSTERS: GENOMICS/PROTEOMICS

**SPONSOR:** ENAR

## 11a. A Nonnegative Matrix Factorization Method for Rank Normalized Data

Danielle Demateis\* and Michael F. Ochs, The College of New Jersey

## 11b. Semiparametric Method in RNA-Seq Differential Expression Analysis Incorporating Uncertainty of Abundance Estimates

Anqi Zhu\*, Joseph G. Ibrahim and Michael I. Love, University of North Carolina, Chapel Hill

## 11c. A Rapid Stepwise Maximum Likelihood Procedure for an Isolation-with-Migration Model

Jieun Park\*, Auburn University; Yujin Chung, Kyonggi University, South Korea

## 11d. Kernel Association Test for Rare Copy Number Variants using Profile Curves

Amanda Brucker\*, Wenbin Lu and Rachel Marceau West, North Carolina State University; Jin Szatkiewicz, University of North Carolina, Chapel Hill; Jung-Ying Tzeng, North Carolina State University

## 11e. Sparse Negative Binomial Model-Based Clustering for RNA-seq Count Data

Md Tanbin Rahman\*, University of Pittsburgh; Tianzhou Ma, University of Maryland; George Tseng, University of Pittsburgh

## 11f. A Hierarchical Bayes Model for Background Correction of Protein Microarrays

Sophie Berube\* and Thomas A. Louis, Johns Hopkins University Bloomberg School of Public Health

## 11g. Statistical Inference of High-Dimensional Modified Poisson-Type Graphical Models with Application to Childhood Asthma in Puerto Ricans

Rong Zhang\* and Zhao Ren, University of Pittsburgh; Wei Chen, Children's Hospital of Pittsburgh of UPMC, University of Pittsburgh

## 11h. Sparse Semiparametric Canonical Correlation Analysis for Data of Mixed Types

Grace Yoon\*, Raymond J. Carroll and Irina Gaynanova, Texas A&M University

## 12. POSTERS: FUNCTIONAL DATA/HIGH DIMENSIONAL

**SPONSOR:** ENAR

## 12a. Computational Methods for Dynamic Prediction

Andrada E. Ivanescu\*, Montclair State University; William Checkley and Ciprian M. Crainiceanu, Johns Hopkins University

## 12b. Modeling Continuous Glucose Monitoring (CGM) Data During Sleep

Irina Gaynanova\*, Texas A&M University; Naresh M. Punjabi and Ciprian M. Crainiceanu, Johns Hopkins University

## 12c. Sampling Studies for Longitudinal Functional Data Analysis

Toni L. Jassel\* and Andrada Ivanescu, Montclair State University

## 12d. Wearable Devices are Objective but Imperfect - Towards Correcting for Two Sources of Error

Dane R. Van Domelen\* and Vadim Zippunnikov, Johns Hopkins University

## 13. POSTERS: SPATIAL/TEMPORAL MODELING

**SPONSOR:** ENAR

## 13a. Spatial Statistical Methods to Assess the Relationship Between Water Violations and Poverty at the County Level: In America, Who has Access to Clean Water?

Ruby Lee Bayliss\* and Loni Phillip Tabb, Drexel University

## 13b. Describing the Spatiotemporal Patterning of Overall Health in the United States using County Health Rankings from 2010-2018

Angel Gabriel Ortiz\* and Loni Tabb, Drexel University

## 13c. Discriminant Analysis for Longitudinal MRI

Rejaul Karim\* and Taps Maiti, Michigan State University; Chae Young Lim, Seoul National University

## 13d. Propensity Score Matching for Multi-Level and Spatial Data

Behzad Kianian\*, Howard H. Chang, Rachel E. Patzer and Lance A. Waller, Emory University

# SCIENTIFIC PROGRAM

SUNDAY, MARCH 24

## 13e. Multiple Testing and Estimation of Disease Associations Based on Semi-Parametric Hierarchical Mixture Models, Possibly Incorporating Brain Areas

Ryo Emoto\*, Takahiro Otani and Shigeyuki Matsui, Nagoya University School of Medicine

## 13f. Trends in Tract-Level Dental Visit Rates in Philadelphia by Race, Space and Time

Guangzi Song\* and Harrison S. Quick, Drexel University

## 14. POSTERS: PUBLIC HEALTH/SURVEYS/EHR

*SPONSOR:* ENAR

## 14a. Supervised Dimension Reduction using Bayesian Hierarchical Modeling: A Simulation Study and Application to Ambient Air Pollutants

Ray Boaz\*, Andrew Lawson and John Pearce, Medical University of South Carolina

## 14b. Comparison of Interval Estimation Methods for Dose-Response Relationship: Frequentist Model Averaging (FMA) versus Corrected Confidence Interval Estimation (CCI) when Exposure Uncertainty is Complex

Deukwoo Kwon\*, University of Miami

## 14c. Joint Modelling of Binary and Continuous Measurements in Large Health Surveys and its Application to Network Analysis, Frailty, and Mortality in NHANES 1999-2010

Debangana Dey\*, Johns Hopkins Bloomberg School of Public Health; Irina Gaynanova, Texas A&M University; Vadim Zippunikov, Johns Hopkins Bloomberg School of Public Health

## 14d. A Methodology and SAS Macro to Estimate Consumption of Alcohol from Survey Designs: Application to NHANES and NLSY

Elysia A. Garcia\* and Stacia M. DeSantis, University of Texas Health Science Center

## 14e. A Privacy-Preserving and Communication Efficient Distributed Algorithm for Logistic Regression with Extremely Rare Outcomes or Exposures

Rui Duan\*, Mary Regina Boland, Jason H. Moore and Yong Chen, University of Pennsylvania

## 14f. Bayesian Methods for Estimating the Population Attributable Risk in the Presence of Exposure Misclassification

Benedict Wong\*, Molin Wang and Lorenzo Trippa, Harvard T.H. Chan School of Public Health; Donna Spiegelman, Yale School of Public Health

## 14g. Evaluating the Effects of Attitudes on Health-Seeking Behavior among a Network of People who Inject Drugs

Ashley Buchanan\*, University of Rhode Island; Ayako Shimada, Thomas Jefferson University; Natallia Katenka, University of Rhode Island; Samuel Friedman, National Development and Research Institutes, Inc.



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

MONDAY, MARCH 25

8:30—10:15 A.M.

## 15. NOVEL NEUROIMAGING METHODS FROM PROCESSING TO ANALYSIS

Franklin Hall 13, 4th Floor

**SPONSOR:** ASA Biometrics Section, ASA Statistics in Imaging Section

**ORGANIZERS:** Inna Chervoneva, Thomas Jefferson University and Kristin A. Linn, University of Pennsylvania

**CHAIR:** Inna Chervoneva, Thomas Jefferson University

### 8:30 A Unified Framework for Brain Functional Connectivity Using Covariance Regression

Ani Eloyan\*, Brown University

### 9:00 Robust Spatial Extent Inference with a Semiparametric Bootstrap Joint Testing Procedure

Simon Vandekar\*, Vanderbilt University; Theodore D. Satterthwaite, Cedric H. Xia, Kosha Ruparel, Ruben C. Gur, Raquel E. Gur and Russell T. Shionhara, University of Pennsylvania

### 9:30 Addressing Partial Volume Effects Using Intra-Subject Locally Adjusted Cerebral Blood Flow Images

Kristin A. Linn\*, University of Pennsylvania; Simon Vandekar, Vanderbilt University; Russell T. Shionhara, University of Pennsylvania

### 10:00 Discussant: Nicole Lazar, University of Georgia

## 16. STATISTICAL CHALLENGES AND OPPORTUNITIES FOR ANALYSIS OF LARGE-SCALE OMICS DATA

Salon B, 5th Floor

**SPONSOR:** ASA Statistics in Genomics and Genetics Section, ASA Statistics in Imaging Section, ASA Statistical Learning and Data Science Section, ASA Mental Health Statistics Section

**ORGANIZER:** Lingzhou Xue, The Pennsylvania State University

**CHAIR:** Yuan Huang, University of Iowa

### 8:30 Linear Hypothesis Testing for High Dimensional Generalized Linear Models

Runze Li\*, The Pennsylvania State University; Chengchun Shi and Rui Song, North Carolina State University; Zhao Chen, Fudan University

### 8:55 Scalable Whole Genome Sequencing Association Analysis using Functional Annotation and Cloud Computing

Xihong Lin\*, Harvard University

### 9:20 Learning Hierarchical Interactions in High Dimensions

Lingzhou Xue\*, The Pennsylvania State University

### 9:45 Analysis of Imaging Genetic “Big Data Squared” Studies

Heping Zhang\*, Yale University School of Public Health

### 10:10 Floor Discussion

## 17. LONGITUDINAL AND FUNCTIONAL MODELS FOR PREDICTING CLINICAL OUTCOMES

Salon C, 5th Floor

**SPONSOR:** ENAR, ASA Bayesian Statistical Science Section, ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Statistics in Imaging Section

**ORGANIZER:** Abdus Sattar, Case Western Reserve University School of Medicine

**CHAIR:** Seunghee Margevicius, Case Western Reserve University School of Medicine

### 8:30 Bayesian Regression Models for Big Spatially or Longitudinally Correlated Functional Data

Jeffrey S. Morris\*, University of Texas MD Anderson Cancer Center; Hongxiao Zhu, Virginia Tech University; Hojin Yang and Michelle Miranda, University of Texas MD Anderson Cancer Center; Wonyul Lee, U.S. Food and Drug Administration; Veera Baladandayuthapani, University of Michigan; Birgir Hrafnkelsson, University of Iceland

### 9:00 Joint Modeling of Multivariate Longitudinal Data with a Binary Response

Paul S. Albert\* and Sung Duk Kim, National Cancer Institute, National Institutes of Health

### 9:30 Modeling of High-Dimensional Clinical Longitudinal Oxygenation Data

Abdus Sattar\* and Seunghee Margevicius, Case Western Reserve University

### 10:00 Discussant:

Dipak Dey, University of Connecticut

## 18. RECENT BAYESIAN METHODS FOR CAUSAL INFERENCE

Salon D, 5th Floor

**SPONSOR:** ASA Bayesian Statistical Science Section, ASA Biometrics Section

**ORGANIZER:** Chanmin Kim, Boston University School of Public Health

**CHAIR:** Chanmin Kim, Boston University School of Public Health

### 8:30 Assessing Causal Effects in the Presence of Treatment Switching through Principal Stratification

Fabrizia Mealli\*, University of Florence

### 8:55 Bayesian Nonparametric Models with Faster Algorithms for Estimating Causal Effects

Jason Roy\*, Rutgers University

### 9:20 Reciprocal Graphical Models for Integrative Gene Regulatory Network Analysis

Peter Mueller\*, University of Texas, Austin; Yang Ni, Texas A&M University; Yuan Ji, University of Chicago

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 9:45 A Bayesian Semiparametric Framework for Causal Inference in High-Dimensional Data

Joseph Antonelli\*, University of Florida; Francesca Dominici, Harvard T.H. Chan School of Public Health

## 10:10 Floor Discussion

### 19. NEW METHODS FOR COST-EFFECTIVENESS ANALYSIS IN HEALTH POLICY RESEARCH

Franklin Hall 3, 4th Floor

**SPONSOR:** ENAR, ASA Health Policy Statistics Section

**ORGANIZER:** Nandita Mitra, University of Pennsylvania

**CHAIR:** Nandita Mitra, University of Pennsylvania

## 8:30 Cost and Cost-Effectiveness Analysis: Where are we now?

Heejung Bang\*, University of California, Davis

## 8:55 Agent-Based Modelling for Better Understanding Health Disparities

Efrén Cruz Cortés\* and Debashis Ghosh, Colorado School of Public Health

## 9:20 Approaches to Cost-Effectiveness Analysis based on Subject-Specific Monetary Value

Andrew J. Spieker\*, Vanderbilt University Medical Center; Nicholas Illenberger, University of Pennsylvania Perelman School of Medicine; Jason A. Roy, Rutgers School of Public Health; Nandita Mitra, University of Pennsylvania Perelman School of Medicine

## 9:45 Microsimulations for Cost-Effectiveness Analysis: Modeling Therapy Sequence in Advanced Cancer

Elizabeth A. Handorf\*, Fox Chase Cancer Center; Andres Correa, Cooper University Health Care; Chethan Ramamurthy, University of Texas Health Science Center at San Antonio; Daniel Geynisman and J. Robert Beck, Fox Chase Cancer Center

## 10:10 Floor Discussion

### 20. FOUNDATIONS OF STATISTICAL INFERENCE IN THE ERA OF MACHINE LEARNING

Franklin Hall 2, 4th Floor

**SPONSOR:** IMS

**ORGANIZER:** Yifan Cui, University of Pennsylvania

**CHAIR:** Michael Kosorok, University of North Carolina, Chapel Hill

## 8:30 Deep Fiducial Inference

Jan Hannig\* and Gang Li, University of North Carolina, Chapel Hill

## 9:00 Uncertainty Quantification of Treatment Regime in Precision Medicine by Confidence Distributions

Minge Xie\*, Yilei Zhan and Sijian Wang, Rutgers University

## 9:30 Healthier Fast Food: Automated Debiasing of Bayesian Posteriors

Keli Liu\*, Stanford University; Xiao-Li Meng, Harvard University

## 10:00 Discussant:

Jamie Robins, Harvard University

### 21. CONTRIBUTED PAPERS: CLINICAL TRIALS: CANCER APPLICATIONS AND SURVIVAL ANALYSIS

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Varadan V. Sevilimedu, Yale University School of Public Health

## 8:30 Comparison of Population Registry Observational Studies and Randomized Clinical Trials in Oncology

Holly E. Hartman\*, Payal D. Soni, Robert T. Dess, Ahmed Abugharib and Steven G. Allen, University of Michigan; Felix Y. Feng, University of California, San Francisco; Anthony L. Zietman, Massachusetts General Hospital; Reshma Jaggi, Daniel E. Spratt and Matthew J. Schipper, University of Michigan

## 8:45 Statistical Considerations for Trials that Study Subpopulation Heterogeneity

Alexander M. Kaizer\*, University of Colorado Denver; Brian P. Hobbs, Cleveland Clinic; Nan Chen, University of Texas MD Anderson Cancer Center; Joseph S. Koopmeiners, University of Minnesota

## 9:00 Designing Clinical Trials with Restricted Mean Survival Time Endpoint as Practical Considerations

Anne Eaton\*, University of Minnesota; Terry Therneau and Jennifer Le-Rademacher, Mayo Clinic

## 9:15 A Generalized Permutation Procedure to Associate Pathways with Clinical Outcomes

Stanley B. Pounds\*, St. Jude Children's Research Hospital; Xueyuan Cao, University of Tennessee Health Science Center

## 9:30 Bayesian Clinical Trial Design for Joint Models of Longitudinal and Survival Data

Jiawei Xu\*, Matthew A. Psioda and Joseph G. Ibrahim, University of North Carolina, Chapel Hill

## 9:45 Interpretation of Time-to-Event Outcomes in Randomized Trials

Ludovic Trinquart\* and Isabelle Weir, Boston University School of Public Health

## 10:00 Floor Discussion

### 22. CONTRIBUTED PAPERS: MULTIPLE TESTING

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jiangtao Gou, Fox Chase Cancer Center

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

**8:30 A Bottom-up Approach to Testing Hypotheses that have a Branching Tree Dependence Structure, with False Discovery Rate Control**

Yunxiao Li\* and Yijuan Hu, Emory University; Glen A. Satten, Centers for Disease Control and Prevention

**8:45 A Global Hypothesis Test for Dependent Endpoints based on Researcher's Predictions**

Robert N. Montgomery\* and Jonathan Mahnken, University of Kansas Medical Center

**9:00 Per-Family Error Rate Control for Gaussian Graphical Model via Knockoffs**

Siliang Gong\*, Qi Long and Weijie Su, University of Pennsylvania

**9:15 Permutations Unlock Experiment-Wise Null Distributions for Large Scale Simultaneous Inference of Microbiome Data**

Stijn Hawinkel\*, Ghent University, Belgium; Luc Bijmens, Janssen Pharmaceutical companies of Johnson and Johnson, Belgium; Olivier Thas, Ghent University, Belgium

**9:30 Global and Simultaneous Hypothesis Testing for High-Dimensional Logistic Regression Models**

Rong Ma\*, Tony T. Cai and Hongzhe Li, University of Pennsylvania

**9:45 Identifying Relevant Covariates in RNA-seq Analysis by Pseudo-Variable Augmentation**

Yet Nguyen\*, Old Dominion University; Dan Nettleton, Iowa State University

**10:00 Asymptotic Simultaneous Confidence Intervals of Odds Ratio in Many-to-One Comparison of Proportions for Correlated Paired Binary Data**

Xuan Peng\* and Chang-Xing Ma, State University of New York at Buffalo

## 23. CONTRIBUTED PAPERS: CLUSTERED DATA METHODS

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Hillary Koch, The Pennsylvania State University

**8:30 Sample Size Estimation for Stratified Individual and Cluster Randomized Trials with Binary Outcomes**

Lee Kennedy-Shaffer\*, Harvard University; Michael D. Hughes, Harvard T.H. Chan School of Public Health

**8:45 Sample Size Considerations for Stratified Cluster Randomization Design with Binary Outcomes and Varying Cluster Size**

Xiaohan Xu\*, University of Texas Southwestern Medical Center and Southern Methodist University; Hong Zhu and Chul Ahn, University of Texas Southwestern Medical Center

**9:00 Power Calculation for Stepped Wedge Designs with Binary Outcomes**

Xin Zhou\*, Harvard T. H. Chan School of Public Health; Xiaomei Liao, AbbVie Inc.; Donna Spiegelman, Yale School of Public Health

**9:15 Pan-Disease Clustering Analysis of the Trend of Period Prevalence**

Chenjin Ma\*, Renmin University of China; Sneha Jadhav, Yale University; Ben-Chang Shia, Taipei Medical University; Shuangge Ma, Yale University

**9:30 Ordinal Clustered Data with Informative Cluster Size in a Longitudinal Study**

Aya A. Mitani\*, Elizabeth K. Kaye and Kerrie P. Nelson, Boston University

**9:45 Inferential Procedures for Consensus Clustering- An ANOVA Based Approach**

Kenneth T. Locke, Jr.\*, Yong Chen and J. Richard Landis, University of Pennsylvania

**10:00 Informatively Empty Clusters with Application to Transgenerational Studies**

Glen W. McGee\* and Marc G. Weisskopf, Harvard University; Marianthi-Anna Kioumourtoglou, Columbia University; Brent A. Coull and Sebastien Haneuse, Harvard University

## 24. CONTRIBUTED PAPERS: GENOME WIDE ASSOCIATION STUDIES AND OTHER GENETIC STUDIES

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Misrak Gezmu, National Institute of Allergy and Infectious Diseases, National Institutes of Health

**8:30 G-SMUT: Generalized Multi-SNP Mediation Intersection-Union Test**

Wujuan Zhong\*, Cassandra N. Spracklen, Karen L. Mohlke, Xiaojing Zheng, Jason Fine and Yun Li, University of North Carolina, Chapel Hill

**8:45 A Multiple-Weighted False Discovery Rate Controlling Procedure in Genome-Wide Association Studies**

Zhou Fang\*, Brigham and Women's Hospital; Nikolaos Patsopoulos, Brigham and Women's Hospital and Broad Institute

**9:00 Phenome-Wide SNP-Set Association Test Based on GWAS Summary Data to Identify Novel Disease-Gene Association**

Bin Guo\* and Baolin Wu, University of Minnesota

**9:15 Polygenic Risk Prediction using Functional Annotation: Application to the International Lung Cancer Consortium (ILCCO)**

Jingwen Zhang\* and Xihong Lin, Harvard T.H. Chan School of Public Health

**9:30 A Simple and General Colocalization Test**

Yangqing Deng\* and Wei Pan, University of Minnesota

Denotes student award winner



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 9:45 Integrative Gene-Based Association Testing for Cancer Phenotypes with Somatic Tumor Expression Data and GWAS Summary Data

Jack W. Pattee\* and Wei Pan, University of Minnesota

## 10:00 Floor Discussion

### 25. CONTRIBUTED PAPERS: TIME SERIES

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jun Young Park, University of Minnesota

## 8:30 Empirical Localized Time-Frequency Analysis via Penalized Reduced Rank Regression

Marie Tuft\* and Robert Todd Krafty, University of Pittsburgh

## 8:45 Multi-Subject Spectral Analysis of Resting-State EEG Signals from Twins Using a Nested Bernstein Dirichlet Prior

Brian B. Hart\*, University of Minnesota; Michele Guindani, University of California, Irvine; Stephen Malone and Mark Fiecas, University of Minnesota

## 9:00 Empirical Frequency Band Analysis of Nonstationary Time Series

Scott A. Bruce\*, George Mason University; Cheng Yong Tang, Temple University; Martica H. Hall and Robert T. Krafty, University of Pittsburgh

## 9:15 Joint Structural Break Detection and Parameter Estimation in High-Dimensional Non-Stationary VAR Models

Abolfazl Safikhani\*, Columbia University; Ali Shojaie, University of Washington, Seattle

## 9:30 Order Restricted Inference in Chronobiology

Yolanda Larriba\*, Cristina Rueda and Miguel A. Fernández, University of Valladolid, Spain; Shyamal D. Peddada, University of Pittsburgh

## 9:45 Dynamic Bayesian Prediction and Calibration using Multivariate Sensor Data Streams

Zhenke Wu\*, Timothy NeCamp and Srijan Sen, University of Michigan

## 10:00 Floor Discussion

MONDAY, MARCH 25

10:15–10:30 P.M.

## REFRESHMENT BREAK WITH OUR EXHIBITORS

Franklin Hall Foyer, 4th Floor

MONDAY, MARCH 25

10:30 A.M.—12:15 P.M.

## 26. DOUGLAS ALTMAN: A CONSUMMATE MEDICAL STATISTICIAN

Franklin Hall 2, 4th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistical Consulting Section, ASA Statistics in Epidemiology Section

**ORGANIZER:** Susan Ellenberg, University of Pennsylvania

**CHAIR:** Kay Dickersin, Johns Hopkins University

## 10:30 What Makes a Great Medical Statistician? The Model of Doug Altman

Steven Goodman\*, Stanford University

## 11:00 Remembering Professor Doug Altman

David Moher\*, Ottawa Hospital Research Institute

## 11:30 Douglas Altman and His Mentoring Legacy

Tianjing Li\*, Johns Hopkins Bloomberg School of Public Health

## 12:00 Discussant:

Cynthia Mulrow, Annals of Internal Medicine

## 27. STATISTICAL ADVANCES FOR EMERGING ISSUES IN HUMAN MICROBIOME RESEARCHES

Salon B, 5th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Jung-Ying Tzeng, North Carolina State University

**CHAIR:** Jung-Ying Tzeng, North Carolina State University

## 10:30 Optimal Permutation Recovery and Estimation of Bacterial Growth Dynamics

Hongzhe Li\* and Yuan Gao, University of Pennsylvania

## 10:55 Higher Criticism Goodness-of-Fit Tests in Phylogenetic Trees for Microbiome Sequencing Experiments

Jeffrey C. Miecnikowski\* and Jiefei Wang, State University of New York at Buffalo (SUNY)

## 11:20 Analyzing Matched Sets of Microbiome Data

Yi-Juan Hu\*, Emory University; Glen A. Satten, Centers for Disease Control and Prevention; Zhengyi Zhu, Emory University

## 11:45 Testing Statistical Interactions Between Microbiome Community Profiles and Covariates

Michael C. Wu\*, Fred Hutchinson Cancer Research Center

## 12:10 Floor Discussion

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 28. WEARABLE TECHNOLOGY IN LARGE OBSERVATIONAL STUDIES

*Franklin Hall 13, 4th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section

**ORGANIZER:** Elizabeth McGuffey, United States Naval Academy

**CHAIR:** Ekaterina Smirnova, Virginia Commonwealth University

### 10:30 Potential Batch Effects and Biases in the UK Biobank Accelerometer Data

John Muschelli\*, Johns Hopkins University

### 10:55 Functional and Compositional Approaches for Accelerometry with Application to the Women's Health Initiative

Chongzhi Di\*, Fred Hutchinson Cancer Research Center

### 11:20 The Tensor Mixture Model for Compositional Data with Essential Zeros: Jointly Profiling Accelerometry-Assessed Physical Activity and Sedentary Behavior in the Hispanic Community Health Study / Study of Latinos (HCHS/SOL)

Daniela Sotres-Alvarez\*, Angel D. Davalos, Jianwen Cai and Kelly Evenson, University of North Carolina, Chapel Hill; Amy H. Herring, Duke University

### 11:45 Functional Regression on Accelerometry Data in the National Health and Nutrition Examination Survey (NHANES)

Elizabeth J. McGuffey\*, United States Naval Academy; Ekaterina Smirnova, Virginia Commonwealth University; Andrew Leroux, Johns Hopkins University; Elham Mokhtari, University of Montana; Vadim Zipunnikov and Ciprian Crainiceanu, Johns Hopkins University

## 12:10 Floor Discussion

## 29. CAUSAL INFERENCE WITH NON-IGNORABLE MISSING DATA: NEW DEVELOPMENTS IN IDENTIFICATION AND ESTIMATION

*Salon C, 5th Floor*

**SPONSORS:** ENAR, ASA Biometrics Section

**ORGANIZER:** Shu Yang, North Carolina State University

**CHAIR:** Mireille Schnitzer, University of Montreal

### 10:30 Identification and Estimation of Causal Effects with Confounders Missing not at Random

Shu Yang\*, North Carolina State University; Linbo Wang, University of Toronto; Peng Ding, University of California, Berkeley

### 10:55 Using Missing Types to Improve Partial Identification with Missing Binary Outcomes

Zhichao Jiang\*, Harvard University; Peng Ding, University of California, Berkeley

### 11:20 Improved Evaluation of HIV Prevalence Adjusting for Informative Non-Participation

Linbo Wang\*, University of Toronto; Eric Tchetgen Tchetgen, University of Pennsylvania; Kathleen Wirth, Harvard T.H. Chan School of Public Health

### 11:45 Bayesian Spatial Propensity Score Analysis: Unmeasured and Geographic Confounding

Joon Jin Song\*, Baylor University; Yawen Guan, North Carolina State University; Veronica Berrocal, University of Michigan; Bo Li, University of Illinois; Shu Yang, North Carolina State University

## 12:10 Floor Discussion

## 30. ADVANCED DEVELOPMENT IN JOINT MODELING AND RISK PREDICTION

*Salon D, 5th Floor*

**SPONSOR:** ENAR

**ORGANIZER:** Ming Wang, The Pennsylvania State University

**CHAIR:** Chixiang Chen, The Pennsylvania State University

### 10:30 Dynamic Prediction of Competing Risk Events using Landmark Sub-Distribution Hazard Model with Multivariate Longitudinal Biomarkers

Liang Li\*, University of Texas MD Anderson Cancer Center

### 10:55 Estimation under Covariate-Induced Dependent Truncation through Inverse Probability of Truncation Weighting

Jing Qian\*, University of Massachusetts; Bella Vakulenko-Lagun, Harvard T.H. Chan School of Public Health; Sy Han Chiou, University of Texas, Dallas; Rebecca A. Betensky, New York University

### 11:20 Joint Modeling of Multiple Time-to-Event Outcomes

Shanshan Zhao\*, National Institute of Environmental Health Sciences, National Institutes of Health; Ross L. Prentice, Fred Hutchinson Cancer Research Center

### 11:45 Predictive Accuracy of Survival Regression Models Subject to Non-Independent Censoring

Ming Wang\*, The Pennsylvania State University; Qi Long, University of Pennsylvania; Chixiang Chen, The Pennsylvania State University

## 12:00 Floor Discussion

## 31. CLASSIFICATION AND VARIABLE SELECTION UNDER ASYMMETRIC LOSS

*Franklin Hall 3, 4th Floor*

**SPONSOR:** IMS

**ORGANIZER:** Xin Tong, University of Southern California

**CHAIR:** Yang Feng, Columbia University

### 10:30 An Umbrella Algorithm to Neyman-Pearson Classification

Xin Tong\*, University of Southern California; Jingyi Jessica Li, University of California, Los Angeles; Yang Feng, Columbia University

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

**10:55 An Umbrella Algorithm that Links Cost-Sensitive Learning to Neyman-Pearson Classification**

Wei Vivian Li\*, University of California, Los Angeles; Xin Tong, University of Southern California; Jingyi Jessica Li, University of California, Los Angeles

**11:20 Neyman-Pearson Classification under Label Noise**

Bradley Rava\* and Shunan Yao, University of Southern California

**11:45 Neyman-Pearson Criterion (NPC): A Budget Constrained Model Selection Criterion for Asymmetric Prediction**

Jingyi Jessica Li\*, University of California, Los Angeles

**12:10 Floor Discussion**

## 32. SPEED POSTERS: HIGH-DIMENSIONAL DATA/OMICS

Franklin Hall 1, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Kristin Linn, University of Pennsylvania

**32a. INVITED SPEED POSTER: Better Diagnostic and Prognostic Tools for Multiple Sclerosis based on MRI**

Russell T. Shinohara\*, University of Pennsylvania

**32b. INVITED SPEED POSTER: Group and Individual Non-Gaussian Component Analysis for Multi-Subject fMRI**

Benjamin B. Risk\*, Emory University; Yuxuan Zhao and David S. Matteson, Cornell University

**32c. Regularized Prediction Modeling in Small Samples with Application to Predicting Toxicity in a CAR T-Cell Immunotherapy Trial**

Mackenzie J. Edmondson\*, University of Pennsylvania; David T. Teachey, Children's Hospital of Philadelphia; Pamela A. Shaw, University of Pennsylvania

**32d. Bayesian GWAS with Structured and Non-Local Priors**

Adam Kaplan\*, Eric F. Lock and Mark Fiecas, University of Minnesota

**32e. Interpretable Advance-Learning for Deriving Optimal Dynamic Treatment Regimes with Observational Data**

Aaron M. Sonabend\* and Tianxi Cai, Harvard T.H. Chan School of Public Health; Peter Szolovits, MIT Computer Science and Artificial Intelligence Laboratory

**32f. Omnibus Weighting Incorporating Multiple Functional Annotations for Whole Genome Sequencing Rare Variant Association Studies**

Xihao Li\*, Zilin Li, Hufeng Zhou and Yaowu Liu, Harvard University; Han Chen, Alanna C. Morrison and Eric Boerwinkle, University of Texas School of Public Health at Houston; Xihong Lin, Harvard University

**32g. Detecting Genes with Abnormal Correlation Among Methylation Sites**

Hongyan Xu\*, Augusta University

**32h. Letting the LaxKAT out of the Bag: Packaging, Simulation, and Neuroimaging Data Analysis for a Powerful Kernel Test**

Jeremy S. Rubin\*, University of Maryland, Baltimore County; Simon Vandekar, Vanderbilt University; Lior Rennert, Clemson University; Mackenzie Edmondson and Russell T. Shinohara, University of Pennsylvania

**32i. Functional Data Analysis for Magnetic Resonance Spectroscopy (MRS) Data in Spinocerebellar Ataxias**

Lynn E. Eberly\*, Meng Yao, James Joers and Gulin Oz, University of Minnesota

**32j. A Local Test for Group Differences in Subject-Level Multivariate Density Data**

Jordan D. Dworkin\* and Russell T. Shinohara, University of Pennsylvania

**32k. Semi-Parametric Differential Abundance Analysis for Metabolomics and Proteomics Data**

Yuntong Li\*, Arnold J. Stromberg, Chi Wang and Li Chen, University of Kentucky

**32l. MIXnorm: Normalizing Gene Expression Data from RNA Sequencing of Formalin-Fixed Paraffin-Embedding Samples**

Shen Yin\*, Southern Methodist University and University of Texas Southwestern Medical Center; Xinlei Wang, Southern Methodist University; Gaoxiang Jia, Southern Methodist University and University of Texas Southwestern Medical Center; Yang Xie, University of Texas Southwestern Medical Center

## 33. CONTRIBUTED PAPERS: PERSONALIZED MEDICINE

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Ming Tang, University of Michigan

**10:30 Estimating Individualized Treatment Regimes from Crossover Studies**

Crystal T. Nguyen\*, Daniel J. Lockett, Grace E. Shearrer and Anna R. Kahkoska, University of North Carolina, Chapel Hill; Donna Spruijt-Metz, University of Southern California; Jaimie N. Davis, University of Texas, Austin; Michael R. Kosorok, University of North Carolina, Chapel Hill

**10:45 Biomarker Screening in the Learning of Individualized Treatment Rules via Net Benefit Index**

Yiwan Zhou\*, University of Michigan; Haoda Fu, Eli Lilly and Company; Peter X.K. Song, University of Michigan

**11:00 Integrative Learning to Combine Individualized Treatment Rules from Multiple Randomized Trials**

Xin Qiu\*, Janssen Research & Development; Donglin Zeng, University of North Carolina, Chapel Hill; Yuanjia Wang, Columbia University

Denotes student award winner



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 11:15 Disparity Subtyping: Bringing Precision Medicine Closer to Disparity Science

Huilin Yu\* and J. Sunil Rao, University of Miami; Jean Eudes Dazard, Case Western Reserve University

## 11:30 Modified Thompson Sampling for Precision Medicine

John Sperger\* and Michael R. Kosorok, University of North Carolina, Chapel Hill

## 11:45 Estimating Individualized Decision Rules with Tail Controls

Zhengling Qi\*, University of North Carolina, Chapel Hill; Jong-Shi Pang, University of Southern California; Yufeng Liu, University of North Carolina, Chapel Hill

## 12:00 Floor Discussion

### 34. CONTRIBUTED PAPERS: EPIDEMIOLOGIC METHODS

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Frank Li, Duke University

## 10:30 Bayesian Piecewise Linear Mixed Models with a Random Change Point: An Application to Study Early Growth Patterns in the Development of Type 1 Diabetes

Xiang Liu\*, Yangxin Huang, Kendra Vehik, Jeffrey Krischer, The TEDDY Study Group, University of South Florida

## 10:45 A Spatial Bayesian Hierarchical Model for Combining Data from Passive and Active Infectious Disease Surveillance Systems

Xintong Li\*, Howard Chang and Lance Waller, Emory University; Qu Cheng, Philip Collender and Justin Remais, University of California, Berkeley

## 11:00 A Joint Model of Opioid Treatment Admissions and Deaths for Adults and Adolescents in Ohio Counties

David M. Kline\*, The Ohio State University; Staci A. Hepler, Wake Forest University

## 11:15 Inference for Case-Control Studies including Prevalent Cases, and Prospective Survival Information

Soutrik Mandal\*, National Cancer Institute, National Institutes of Health; Jing Qin, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Ruth Pfeiffer, National Cancer Institute, National Institutes of Health

## 11:30 Effect Size Measures for Mediation Analysis of Multiple Correlated Exposures

Yue Jiang\*, University of North Carolina, Chapel Hill; Shanshan Zhao, National Institute of Environmental Health Sciences, National Institutes of Health; Jason Peter Fine, University of North Carolina, Chapel Hill

## 11:45 Regression Analysis of Combined Incident and Prevalent Cohort Data

Chi Hyun Lee\*, University of Massachusetts; Jing Ning, University of Texas MD Anderson Cancer Center; Richard Kryscio, University of Kentucky; Yu Shen, University of Texas MD Anderson Cancer Center

## 12:00 Floor Discussion

### 35. CONTRIBUTED PAPERS: STATISTICAL GENETICS: SINGLE-CELL SEQUENCING/TRANSCRIPTOMIC DATA

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jian Hu, University of Pennsylvania, Perelman School of Medicine

## 10:30 Single-Cell RNA Sequencing: Normalization for Technical Noise and Batch Effects

Nicholas J. Lytal\*, Di Ran and Lingling An, University of Arizona

## 10:45 Bulk Gene Expression Deconvolution by Single-Cell RNA Sequencing

Meichen Dong\*, Yuchao Jiang and Fei Zou, University of North Carolina, Chapel Hill

## 11:00 Bivariate Zero-Inflated Negative Binomial (BZINB) Model for Measuring Dependence

Hunryong Cho\* and Di Wu, University of North Carolina, Chapel Hill

## 11:15 SCOPE: A Normalization and Copy Number Estimation Method for Single-Cell DNA Sequencing

Rujin Wang\*, Danyu Lin and Yuchao Jiang, University of North Carolina, Chapel Hill

## 11:30 MACAM: A Semi-Supervised Statistical Deconvolution Method for Mixed Transcriptomic Data

Li Dong\*, Fei Zou and Xiaojing Zheng, University of North Carolina, Chapel Hill

## 11:45 Detecting Regulatory Genetic Variants with Transcription Factor Binding Affinity Testing

Sunyoung Shin\*, University of Texas, Dallas; Chandler Zuo, A.R.T. Advisors; Sunduz Keles, University of Wisconsin, Madison

## 12:00 Floor Discussion

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 36. CONTRIBUTED PAPERS: MACHINE LEARNING AND TESTING WITH HIGH DIMENSIONAL DATA

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Toyya A. Pujol-Mitchell, Georgia Institute of Technology and Harvard Medical School

- 10:30** **Integrative Linear Discriminant Analysis with Guaranteed Error Rate Improvement**  
Quefeng Li\*, University of North Carolina, Chapel Hill; Lexin Li, University of California, Berkeley

- 10:45** **High-Dimensional Decomposition-Based Canonical Correlation Analysis**  
Hai Shu\*, University of Texas MD Anderson Cancer Center; Xiao Wang, Purdue University; Hongtu Zhu, University of North Carolina, Chapel Hill; Peng Wei, University of Texas MD Anderson Cancer Center

- 11:00** **Estimation of Tumor Immune Cell Content Using Single-Cell RNA-seq Data**  
Christopher M. Wilson\*, Xuefeng Wang and Xiaoqing Yu, H. Lee Moffitt Cancer Center

- 11:15** **Informative Projections and Dimension Reductions for High-Dimensional Data Clustering**  
Zhipeng Wang\*, Genentech; David Scott, Rice University

- 11:30** **Simultaneous Estimation of Number of Clusters and Feature Sparsity in Clustering High-Dimensional Data using Resampling Methods**  
Yujia Li\*, University of Pittsburgh; Xiangrui Zeng, Carnegie Mellon University; Chien-Wei Lin, Medical College of Wisconsin; George Tseng, University of Pittsburgh

- 11:45** **An Evaluation of Machine Learning and Classical Statistical Methods for Discovery in Large-Scale Translational Data**  
Megan C. Hollister\* and Jeffrey D. Blume, Vanderbilt University

- 12:00** **Floor Discussion**

## 37. CONTRIBUTED PAPERS: SPATIO-TEMPORAL MODELING

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Scott A. Bruce, George Mason University

- 10:30** **Spatio-Temporal Models of Intrahepatic Hepatitis C Virus Propagation in Humans**  
Paula Moraga\* and Peter J. Diggle, Lancaster Medical School; Ruy M. Ribeiro, Universidade de Lisboa; Ashwin Balagopal, Johns Hopkins University; Benjamin M. Taylor, Lancaster Medical School

Denotes student award winner

- 10:45** **Simultaneous Ranking and Clustering of Small Areas based on Health Outcomes using Nonparametric Empirical Bayes Methods**  
Ronald E. Gangnon\* and Cora Allen-Coleman, University of Wisconsin, Madison

- 11:00** **Spatial-Temporal Models and Validation for Predicting Historical Missing Cancer Incidence**  
Benmei Liu\*, Li Zhu and Huann-Sheng Chen, National Cancer Institute, National Institutes of Health; Joe Zou, IMS; Rebecca Siegel, Kim D. Miller and Ahmedin Jemal, American Cancer Society; Eric J. Feuer, National Cancer Institute, National Institutes of Health

- 11:15** **Using Spatiotemporal Models to Generate Synthetic Data for Public Use**  
Harrison Quick\*, Drexel University; Lance Waller, Emory University

- 11:30** **Bayesian Selection of Neighborhood Structure in Spatial Model**  
Marie Denis\*, CIRAD; Benoît Cochard, PalmElit

- 11:45** **Floor Discussion**

**MONDAY, MARCH 25**  
**12:15—1:30 P.M.**

## ROUNDTABLE LUNCHEONS

Salon F, 5th Floor

**MONDAY, MARCH 25**  
**1:45—3:30 P.M.**

## 38. EMERGING STATISTICAL ISSUES AND METHODS FOR INTEGRATING MULTI-DOMAIN MHEALTH DATA

Salon B, 5th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Statistical Learning and Data Science Section, ASA Mental Health Statistics Section

**ORGANIZER:** Haochang Shou, University of Pennsylvania

**CHAIR:** Fengqing (Zoe) Zhang, Drexel University

- 1:45** **Methods for Combining Activity Information from Heart Rate and Accelerometry in the Baltimore Longitudinal Study of Aging**  
Ciprian Crainiceanu\*, Johns Hopkins University

- 2:10** **RAR: A Rest-Activity Rhythm Data Analysis Package in R**  
Jessica Graves\*, Haoyi Fu, Robert Krafty, Stephen Smagula and Matricia Hall, University of Pittsburgh

- 2:35** **Mixed Effects Neural Networks for Longitudinal Data**  
Ian J. Barnett\*, University of Pennsylvania

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 3:00 Objective Measurement Versus Performance Test for Physical Activity: Which to Use?

Jiawei Bai\*, Vadim Zipunnikov, Lisa M. Reider and Daniel O. Scharfstein, Johns Hopkins University

## 3:25 Floor Discussion

### 39. CAUSAL INFERENCE WITH DIFFERENCE-IN-DIFFERENCES AND REGRESSION DISCONTINUITY DESIGNS

Franklin Hall 13, 4th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Health Policy Statistics Section

**ORGANIZER:** Fan Li, Duke University

**CHAIR:** Jason Roy, Rutgers University

## 1:45 Patterns of Effects and Sensitivity Analysis for Differences-in-Differences

Luke Keele\*, University of Pennsylvania; Colin Fogarty, Massachusetts Institute of Technology; Jesse Hsu and Dylan Small, University of Pennsylvania

## 2:10 A Bracketing Relationship Between the Difference-in-Differences and Lagged Dependent Variable Adjustment

Fan Li\*, Duke University; Peng Ding, University of California, Berkeley

## 2:35 Augmented Weighting Estimators for Difference-in-Differences

Frank Li\* and Fan Li, Duke University

## 3:00 A Regression Discontinuity Approach for Addressing Temporal Confounding in the Evaluation of Therapeutic Equivalence of Brand and Generic Drugs

Ravi Varadhan\*, Lamar Hunt, Dan Scharfstein, Irene Murimi and Jodi Segal, Johns Hopkins University

## 3:25 Floor Discussion

### 40. STATISTICAL CHALLENGES IN SYNTHESIZING ELECTRONIC HEALTHCARE DATA

Franklin Hall 2, 4th Floor

**SPONSOR:** ENAR, ASA Biopharmaceutical Section

**ORGANIZER:** Junjing Lin, AbbVie

**CHAIR:** Margaret Gamalo-Siebers, Eli Lilly and Company

## 1:45 Comparing Real World Data with Randomized Trial Results to Assess Validity: Preliminary Insights from the RCT DUPLICATE Project

Jessica M. Franklin\* and Sebastian Schneeweiss, Brigham and Women's Hospital and Harvard Medical School

## 2:10 Population-Level Effect Estimation: From Art to Science

Martijn J. Schuemie\*, Janssen R&D

## 2:35 Robust Privacy-Preserving Statistical Methods for Horizontally Partitioned Incomplete Data in Distributed Health Data Networks

Qi Long\* and Changgee Chang, University of Pennsylvania; Yi Deng, Google; Xiaoqian Jiang, University of Texas Health Science Center at Houston

## 3:00 Opportunities and Challenges in Leveraging Real World Data in Regulatory Clinical Studies

Heng Li\* and Lilly Q. Yue, U.S. Food and Drug Administration

## 3:25 Floor Discussion

### 41. USING HISTORICAL DATA TO INFORM DECISIONS IN CLINICAL TRIALS: EVIDENCE BASED APPROACH IN DRUG DEVELOPMENT

Franklin Hall 3, 4th Floor

**SPONSOR:** ASA Biopharmaceutical Section

**ORGANIZER:** Satrajit Roychoudhury, Pfizer Inc.

**CHAIR:** Russell Shinhora, University of Pennsylvania

## 1:45 Leveraging Historical Controls Using Multisource Adaptive Design

Brian P. Hobbs\*, Cleveland Clinic

## 2:15 Use of Historical Data for Premarketing Evaluation of Medical Devices

Nelson Lu\*, Yunling Xu and Lilly Yue, U.S. Food and Drug Administration, Center for Devices and Radiological Health

## 2:45 Evidence Synthesis of Time-to-Event Data in Design and Analysis of Clinical Trials

Satrajit Roychoudhury\*, Pfizer Inc.

## 3:15 Discussant:

Kert Viele, Berry Consultants

### 42. STATISTICAL INNOVATIONS IN SINGLE-CELL GENOMICS

Salon C, 5th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Mingyao Li, University of Pennsylvania

**CHAIR:** Nancy Zhang, University of Pennsylvania

## 1:45 Characterizing Technical Artifacts in Single-Cell RNA-Seq Data Using A Data Generation Simulation Framework

Rhonda Bacher\*, University of Florida; Christina Kendziorski, University of Wisconsin, Madison; Li-Fang Chu and Ron Stewart, Morgridge Institute for Research

## 2:10 Hi-C Deconvolution via Joint Modeling of Bulk and Single-Cell Hi-C Data

Yun Li\*, Ruth Huh, Yuchen Yang and Jin Szatkiewicz, University of North Carolina, Chapel Hill



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 2:35 Semi-Soft Clustering of Single Cell Data

Kathryn Roeder\* and Lingxue Zhu, Carnegie Mellon University; Bernie Devlin, University of Pittsburgh; Jing Lei, Carnegie Mellon University; Lambertus Klei, University of Pittsburgh

## 3:00 Evaluation of Cell Clustering in Single Cell Data

Zhijin (Jean) Wu\*, Brown University

## 3:25 Floor Discussion

### 43. ADVANCES IN STATISTICAL METHODS FOR SURVEILLANCE DATA OF INFECTIOUS DISEASES

Salon D, 5th Floor

**SPONSOR:** IMS

**ORGANIZER:** Yang Yang, University of Florida

**CHAIR:** Eben Kenah, The Ohio State University

## 1:45 Statistical Challenges When Analysing Emerging Epidemic Outbreaks

Tom Britton\*, Stockholm University; Gianpaolo Scalia Tomba, University of Rome Tor Vergata

## 2:10 Efficient Bayesian Semiparametric Modeling and Variable Selection for Spatio-Temporal Transmission of Multiple Pathogens

Nikolay Bliznyuk\*, University of Florida; Xueying Tang, Columbia University

## 2:35 Dynamic Monitoring of Spatio-Temporal Disease Incidence Rates

Peihua Qiu\*, University of Florida

## 3:00 Statistical Adjustment for Reporting Bias in Surveillance Data of Infectious Diseases

Yang Yang\*, Tim Tsang, Diana Rojas Alvarez and Ira Longini, University of Florida; M. Elizabeth Halloran, Fred Hutchinson Cancer Research Center

## 3:25 Floor Discussion

### 44. SPEED POSTERS: SPATIO-TEMPORAL MODELING/ LONGITUDINAL DATA/SURVIVAL ANALYSIS

Franklin Hall 1, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jason Liang, National Institute of Allergy and Infectious Diseases

## 44a. INVITED SPEED POSTER: A Multivariate Spatio-Temporal Model of Opioid Overdose Deaths in Ohio

Staci Hepler\*, Wake Forest University; David Kline, The Ohio State University

## 44b. INVITED SPEED POSTER: A Multivariate Dynamic Spatial Factor Model for Speciated Pollutants and Adverse Birth Outcomes

Kimberly A. Kaufeld\*, Los Alamos National Laboratory

## 44c. Forecasting Cancer Incidence and Mortality under the Age-Period-Cohort Model

Ana F. Best\* and Philip S. Rosenberg, National Cancer Institute, National Institutes of Health

## 44d. A Hierarchical Bayesian Approach to Predicting Time-to-Conversion to Alzheimer's Disease using a Longitudinal Map of Cortical Thickness

Ning Dai\*, University of Minnesota; Hakmook Kang, Vanderbilt University; Galin Jones and Mark Fiecas, University of Minnesota

## 44e. Body Mass Index and Breast Cancer Survival: A Censored Quantile Regression Analysis

Dawen Sui\* and Mary Elizabeth Edgerton, University of Texas MD Anderson Cancer Center

## 44f. Subgroup Analyses using Cox Regression: Relaxing the Proportional Hazards (PH) Assumption

Karen Chiswell\*, Amanda Brucker and Adrian Coles, Duke Clinical Research Institute; Yuliya Lokhnygina and Megan L. Neely, Duke University; Adam Silverstein and Daniel M. Wojdyla, Duke Clinical Research Institute

## 44g. Comparison of Methods for Analyzing Time-Varying Continuous Covariates in Survival Analysis

Qian Liu\* and Abigail R. Smith, Arbor Research Collaborative for Health; Laura H. Mariani and Viji Nair, University of Michigan; Jarcy Zee, Arbor Research Collaborative for Health

## 44h. Joint Modelling of Sequential Time-to-Events

Akhtar Hossain\*, University of South Carolina, Columbia; Hrishikesh Chakraborty, Duke Clinical Research Institute

## 44i. An Alternative to the Cox Model Estimator in Time-to-Event Clinical Trials with Treatment of Physician's Choice

Philani B. Mpofu\*, Indiana University, Richard M. Fairbanks School of Public Health; Stella W. Karuri, U.S. Food and Drug Administration

## 44j. Phase II Trial Design with Growth Modulation Index as the Primary Endpoint

Jianrong John Wu\*, University of Kentucky

## 44k. Multilevel Stochastic Block Model for Dynamic Networks

Jihui Lee\*, Weill Cornell Medicine; Jeff Goldsmith and Gen Li, Columbia University

## 44l. Integrative Variable Selection Method for Subgroup Analyses in Longitudinal Data

Xiaochen Li\* and Sujuan Gao, Indiana University

## 44m. Evaluation of the Performance of Propensity Score Weighting Methods for Survival Outcomes

Uma Siangphoe\* and Kwan R. Lee, Janssen Research & Development

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 45. CONTRIBUTED PAPERS: BIOMARKERS

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Zhengling Qi, University of North Carolina, Chapel Hill

### 1:45 Group Testing Estimation for Multiple Infections with Adjustments for Dilution Effects

Md S. Warasi\*, Radford University; Hrishikesh Chakraborty, Duke University

### 2:00 Nonparametric Conditional Density Estimation for Pooled Biomarker Data

Dewei Wang\*, Xichen Mou and Joshua Tebbs, University of South Carolina

### 2:15 A Mixture of Tukey's g-h Distributions With Application to Measuring Protein Biomarkers

Tingting Zhan\* and Inna Chervoneva, Thomas Jefferson University

### 2:30 Bayesian Nonparametric Clustering Analysis for Multi-Scale Molecular Data

Yize Zhao\*, Weill Cornell Medicine

### 2:45 Floor Discussion

## 46. CONTRIBUTED PAPERS: BAYESIAN MODELING AND VARIABLE SELECTION

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Sharang Chaudhry, University of Nevada Las Vegas

### 1:45 Bayesian Sparse Envelope Model for Multivariate Linear Regression

Minji Lee\*, University of Florida; Saptarshi Chakraborty, Memorial Sloan Kettering Cancer Center; Zhihua Su and Malay Ghosh, University of Florida

### 2:00 Interaction Detection Using Bayesian Decision Tree Ensembles

Junliang Du\* and Antonio R. Linero, Florida State University

### 2:15 Bayesian Selection of Variance Components in Linear Mixed Models

Benjamin Heuclin\*, Université de Montpellier, France; Marie Denis and Frédéric Mortier, CIRAD, France; Catherine Trottier, Université de Montpellier, France

### 2:30 Bayesian Hierarchical Modeling on Covariance Valued Data

Satwik Acharyya\*, Texas A&M University; Zhengwu Zhang, University of Rochester; Anirban Bhattacharya and Debdeep Pati, Texas A&M University

### 2:45 Knowledge-Guided Bayesian Variable Selection in Non-Linear Support Vector Machines for Structured High-Dimensional Data

Wenli Sun\*, Changge Chang and Qi Long, University of Pennsylvania

### 3:00 Bayesian Adjustment for Confounding when Estimating Average Causal Effects for Time-to-Event Outcomes

Li Xu\* and Chi Wang, University of Kentucky

### 3:15 BayesESS: An R package and a Web-based Application for Quantifying the Impact of Parametric Priors in Bayesian Analysis

Jaejoon Song\*, U.S. Food and Drug Administration; Jiun-Kae Jack Lee, University of Texas MD Anderson Cancer Center; and Satoshi Morita, Kyoto University

## 47. CONTRIBUTED PAPERS: FUNCTIONAL DATA APPLICATIONS AND METHODS

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Marzia A. Cremona, The Pennsylvania State University

### 1:45 Activity Classification Using the Smartphone Gyroscope and Accelerometer

Emily Huang\* and Jukka-Pekka Onnela, Harvard T.H. Chan School of Public Health

### 2:00 Quantifying Physical Activity with Smartphone Accelerometry

Josh Barback\*, Harvard T.H. Chan School of Public Health

### 2:15 Addressing Missing Accelerometer Data with Functional Data Analysis (FDA)

Patrick Hilden\*, Jeff Goldsmith, Joseph Schwartz, Kieth Diaz and Ipek Ensari, Columbia University

### 2:30 A Functional Mixed Model for Scalar on Function Regression with Application to a Functional MRI Study

Wanying Ma\*, Bowen Liu and Luo Xiao, North Carolina State University; Martin A. Lindquist, Johns Hopkins Bloomberg School of Public Health

### 2:45 Covariate-Adjusted Region-Referenced Generalized Functional Linear Model for EEG Data

Aaron W. Scheffler\*, Donatello Telesca, Catherine A. Sugar, Shafali Jeste, Abigail Dickinson, Charlotte DiStefano and Damla Senturk, University of California, Los Angeles

### 3:00 Function-on-Scalar Quantile Regression with Application to Mass Spectrometry Proteomics Data

Yusha Liu\* and Meng Li, Rice University; Jeffrey S. Morris, University of Texas MD Anderson Cancer Center

### 3:15 Identification of Problematic Cell Lines from in Vitro Drug Response Data

Farnoosh Abbas Aghababazadeh\* and Brooke L. Fridley, H. Lee Moffitt Cancer Center

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 48. CONTRIBUTED PAPERS: MACHINE LEARNING AND STATISTICAL RELATIONAL LEARNING

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Quefeng Li, University of North Carolina, Chapel Hill

**1:45** **Using Deep Learning for Automated Scoring of Animal Sleep and Wake States in Neuroscience Research and Development**  
Vladimir Svetnik\*, Ting-Chuan Wang and Yuting Xu, Merck & Co.

**2:00** **Machine Learning for Protein Design**  
Yuting Xu\* and Andy Liaw, Merck & Co.

**2:15** **Estimating Optimal Treatment Regimes Using Multivariate Random Forests**  
Boyi Guo\*, University of Alabama at Birmingham; Ruqing Zhu, Hannah D. Holscher, Loretta S. Auvi, Michael E. Welge and Colleen B. Bushell, University of Illinois at Urbana-Champaign; David J. Baer and Janet A. Novotny, Beltsville Human Nutrition Research Center

**2:30** **Machine Learning Algorithms for Partially Supervised Data with Applications in Group Testing**  
Michael R. Stutz\*, Zichen Ma and Joshua M. Tebbs, University of South Carolina

**2:45** **Robust Nonparametric Methods for Difference-in-Differences Designs**  
Toyra A. Pujol\*, Georgia Institute of Technology; Sherri Rose, Harvard Medical School

**3:00** **Determining the Number of Latent Factors in Statistical Multi-Relational Model**  
Chengchun Shi\*, Wenbin Lu and Rui Song, North Carolina State University

**3:15** **Floor Discussion**

## 49. CONTRIBUTED PAPERS: INTERVAL-CENSORED AND MULTIVARIATE SURVIVAL DATA

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Yizeng He, Medical College of Wisconsin

**1:45** **Multivariate Proportional Intensity Model with Random Coefficients for Event Time Data with Application to Process Data from Educational Assessment**  
Hok Kan Ling\*, Jingchen Liu and Zhiliang Ying, Columbia University

**2:00** **Method for Evaluating Longitudinal Follow-Up Frequency: Application to Dementia Research**  
Leah H. Suttner\* and Sharon X. Xie, University of Pennsylvania

**2:15** **Copula-Based Sieve Semiparametric Transformation Model for Bivariate Interval-Censored Data**

Tao Sun\*, Wei Chen and Ying Ding, University of Pittsburgh

**2:30** **Bayesian Regression Analysis of Multivariate Interval-Censored Failure Time Data Under the Normal Frailty Probit Model**

Yifan Zhang\* and Lianming Wang, University of South Carolina

**2:45** **A Proportional Hazards Model for Interval-Censored Data Subject to Instantaneous Failures**

Prabhashi W. Withana Gamage\*, James Madison University; Monica Chaudari, University of North Carolina, Chapel Hill; Christopher S. McMahan, Clemson University; Edwin H. Kim and Michael R. Kosorok, University of North Carolina, Chapel Hill

**3:00** **Floor Discussion**

**MONDAY, MARCH 25**

**3:30—3:45 P.M.**

**REFRESHMENT BREAK WITH OUR EXHIBITORS**

Franklin Hall Foyer, 4th Floor

**MONDAY, MARCH 25**

**3:45—5:30 P.M.**

## 50. UNDERSTANDING THE COMPLEXITY AND INTEGRITY OF CLINICAL TRIAL DATA

Franklin Hall 2, 4th Floor

**SPONSOR:** ENAR, ASA Biometrics Section

**ORGANIZER:** Pamela Shaw, University of Pennsylvania

**CHAIR:** Susan Ellenberg, University of Pennsylvania

**3:45** **Simulating Realistic Clinical Trial Data**  
Naji Younes\*, The George Washington University

**4:15** **The P-value Requires Context and Correct Interpretation in Clinical Trials**  
Rebecca Betensky\*, NYU College of Global Public Health

**4:45** **Detecting Fraudulent Baseline Data in Clinical Trials**  
Michael A. Proschan\*, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Pamela A. Shaw, University of Pennsylvania Perelman School of Medicine

**5:15** **Discussant:**  
Geert Molenberghs, University of Hasselt and Leuven

Denotes student award winner



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 51. REPLICABILITY IN BIG DATA PRECISION MEDICINE

*Franklin Hall 13, 4th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Naim Rashid, University of North Carolina, Chapel Hill

**CHAIR:** Pedro Baldoni, University of North Carolina, Chapel Hill

### 3:45 Reproducibility and Heterogeneity in Meta-Analysis and Replication of Transcriptomic Studies

George Tseng\*, University of Pittsburgh

### 4:10 Current Topics in Multi-Study Learning

Prasad Patil\* and Giovanni Parmigiani, Harvard T.H. Chan School of Public Health and Dana-Farber Cancer Institute

### 4:35 A Statistical Framework for Measuring Replicability and Reproducibility of High-Throughput Data from Multiple Labs

Qunhua Li\*, The Pennsylvania State University; Monia Ranalli, Tor Vergata University, Rome, Italy; Yafei Lyu, University of Pennsylvania

### 5:00 Modeling Between-Study Heterogeneity for Improved Replicability in Gene Signature Selection and Clinical Prediction

Naim Rashid\*, Quefeng Li, Jen Jen Yeh and Joseph Ibrahim, University of North Carolina, Chapel Hill

### 5:25 Floor Discussion

## 52. COMPUTATIONALLY-INTENSIVE BAYESIAN TECHNIQUES FOR BIOMEDICAL DATA: RECENT ADVANCES

*Salon B, 5th Floor*

**SPONSOR:** ENAR, ASA Bayesian Statistical Science Section, ASA Biometrics Section, ASA Statistics in Epidemiology Section

**ORGANIZER:** Dipankar Bandyopadhyay, Virginia Commonwealth University

**CHAIR:** Bret Zeldow, Harvard University

### 3:45 Finding and Leveraging Structure with Bayesian Decision Tree Ensembles

Antonio R. Linero\* and Junliang Du, Florida State University

### 4:10 Bayesian Estimation of Individualized Treatment-Response Curves in Populations with Heterogeneous Treatment Effects

Yanxun Xu\*, Johns Hopkins University; Yanbo Xu, Georgia Tech University; Suchi Saria, Johns Hopkins University

### 4:35 Monotone Single-Index Models for Highly Skewed Response

Debajyoti Sinha\*, Florida State University; Kumaresh Dhara, University of Florida; Bradley Hupf and Greg Hajcak, Florida State University

### 5:00 A Graphical Model for Skewed Matrix-Variate Non-Randomly Missing Data

Dipankar Bandyopadhyay\*, Virginia Commonwealth University; Lin Zhang, University of Minnesota

### 5:25 Floor Discussion

## 53. MULTIVARIATE FUNCTIONAL DATA ANALYSIS WITH MEDICAL APPLICATIONS

*Salon C, 5th Floor*

**SPONSOR:** ENAR, ASA Statistics in Imaging Section, ASA Statistical Learning and Data Science Section, ASA Mental Health Statistics Section

**ORGANIZER:** Kuang-Yao Lee, Temple University

**CHAIR:** Kuang-Yao Lee, Temple University

### 3:45 Dimension Reduction for Functional Data based on Weak Conditional Moments

Bing Li\*, The Pennsylvania State University; Jun Song, University of North Carolina, Charlotte

### 4:10 Alignment of fMRI Time-Series and Functional Connectivity

Jane-Ling Wang\* and Chun-Jui Chen, University of California, Davis

### 4:35 Functional Marginal Structural Models for Time-Varying Confounding of Mood Assessments

Haochang Shou\*, University of Pennsylvania

### 5:00 Finding Biomarkers for Childhood Obesity Using Functional Data Analysis

Matthew Reimherr\*, Sara Craig, Kateryna Makova and Francesca Chiaromonte, The Pennsylvania State University; Alice Parodi, Milano di Politecnico; Junli Lin and Ana Kenney, The Pennsylvania State University

### 5:25 Floor Discussion

## 54. METHODS FOR EXAMINING HEALTH EFFECTS OF EXPOSURE TO THE WORLD TRADE CENTER ATTACK AND BUILDING COLLAPSE

*Salon D, 5th Floor*

**SPONSORS:** ENAR, ASA Section: Statistics and the Environment, ASA Statistics in Epidemiology Section

**ORGANIZER:** Charles B. Hall, Albert Einstein College of Medicine

**CHAIR:** Shankar Viswanathan, Albert Einstein College of Medicine

### 3:45 Modeling Comorbid Mental and Medical Outcomes via Latent Class Regression

Yongzhao Shao\*, New York University School of Medicine

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

**4:15 Handgrip Strength of World Trade Center Responders: The Long-Term Role of Re-Experiencing Traumatic Events**  
Sean A. Clouston\* and Peifen Kuan, Stony Brook University; Soumyadeep Mukherjee, Rhode Island College; Roman Kotov, Evelyn J. Bromet and Benjamin J. Luft, Stony Brook University

**4:45 Cancer Latency After Environmental Exposure: A Change Point Approach**  
Charles B. Hall\*, Albert Einstein College of Medicine

**5:15 Discussant:**  
Judith Goldberg, New York University School of Medicine

## 55. REGRESSION, MEDIATION, AND GRAPHICAL MODELING TECHNIQUES FOR MICROBIOME DATA

Franklin Hall 3, 4th Floor

**SPONSOR:** IMS

**ORGANIZER:** Aditya K. Mishra, Flatiron Institute, Simons Foundation

**CHAIR:** Christian L. Mueller, Flatiron Institute, Simons Foundation

**3:45 A Framework for Analysis of Microbiome Data with Respect to Multivariate Clinical Instruments**  
Alexander Alekseyenko\*, Medical University of South Carolina

**4:10 Disentangling Microbial Associations via Latent Variable Graphical Models**  
Zachary D. Kurtz\*, Lodo Therapeutics; Christian L. Mueller, Flat Iron Institute, Simons Foundation

**4:35 Robust Regression with Compositional Covariates**  
Aditya Kumar Mishra\* and Christian L. Mueller, Flatiron Institute, Simons Foundation

**5:00 Mediation Analysis in Investigating the Role of Microbiome in Human Health**  
Lingling An\*, Kyle Carter and Meng Lu, University of Arizona

**5:25 Floor Discussion**

## 56. SPEED POSTERS: EHR DATA, EPIDEMIOLOGY, PERSONALIZED MEDICINE, CLINICAL TRIALS

Franklin Hall 1, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Ana Maria Ortega-Villa, National Institute of Allergy and Infectious Diseases

**56a. INVITED SPEED POSTER: Combining Inverse-Probability Weighting and Multiple Imputation to Adjust for Selection Bias Due to Missing Data in EHR-Based Research**  
Sebastien Haneuse\* and Tanayott Thaweethai, Harvard T.H. Chan School of Public Health

**56b. INVITED SPEED POSTER: Methods to Utilize Longitudinal EHR Data to Investigate Whether Moving to a Different Built Environment Affects Health**  
Jennifer F. Bobb\* and Andrea J. Cook, Kaiser Permanente

**56c. Pragmatic Evaluation of Relative Risk Models in PheWAS Analysis**  
Ya-Chen Lin\*, Vanderbilt University; Siwei Zhang, Lisa Bastarache and Todd Edwards, Vanderbilt University Medical Center; Jill M. Pulley, Vanderbilt University; Joshua C. Denny, Vanderbilt University Medical Center; Hakmook Kang and Yaomin Xu, Vanderbilt University

**56d. Operating Characteristics of Bayesian Joint Benefit-Risk Copula Models**  
Nathan T. James\* and Frank E. Harrell, Jr., Vanderbilt University

**56e. Heterogeneity Assessment of Treatment Effect among Subpopulations in Basket Trials**  
Ryo Sadachi\* and Akihiro Hirakawa, The University of Tokyo, Japan

**56f. Extensive Comparisons of the Interval-Based Phase I Design with 3+3 Design for the Trials with 3 or 4 Dose Levels**  
Jongphil Kim\*, H. Lee Moffitt Cancer Center and University of South Florida

**56g. Combining Evidence from Randomized Clinical Trials Across Outbreaks**  
Natalie E. Dean\*, University of Florida; Victor De Gruttola, Harvard University

**56h. Conditional Quantile Inference with Zero-inflated Outcomes**  
Wodan Ling\*, Bin Cheng, Ying Wei and Ken Cheung, Columbia University

**56i. Nonlinear Mixture Model for Modeling Trajectories of Ordinal Markers in Neurological Disorders**  
Qinxia Wang\*, Ming Sun and Yuanjia Wang, Columbia University

**56j. An Augmented Survival Analysis Method for Interval Censored and Mis-Measured Outcomes**  
Chongliang Luo\* and Yong Chen, University of Pennsylvania

**56k. Transformation of Activity Counts from Multiple Activity Monitoring Devices using Latent Correlation**  
Jordan Johns\*, Vadim Zipunnikov and Ciprian Crainiceanu, Johns Hopkins School of Public Health

**56l. Improved Doubly Robust Estimation in Learning Individualized Treatment Rules**  
Yinghao Pan\*, University of North Carolina, Charlotte; Yingqi Zhao, Fred Hutchinson Cancer Research Center

# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

- 56m. Projecting Clinical Trial Results to Alternative Populations by Interpolation**  
Shuang Li\*, Southern Methodist University; Daniel F. Heitjan, Southern Methodist University and University of Texas Southwestern Medical Center

## 57. CONTRIBUTED PAPERS: AGREEMENT MEASURES AND DIAGNOSTICS

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Li C. Cheung, National Cancer Institute, National Institutes of Health

- 3:45 The Impact of Rater Factors on Ordinal Agreement**  
Kerrie Nelson\* and Aya Mitani, Boston University; Don Edwards, University of South Carolina

- 4:00 Robust Matrix-Based Measures of Agreement Based on L-statistics for Repeated Measures**  
Elahe Tashakor\* and Vernon M. Chinchilli, Pennsylvania State College of Medicine

- 4:15 Net Benefit of a Diagnostic Test to Rule-In or Rule-Out a Clinical Condition**  
Gene A. Pennello\*, U.S. Food and Drug Administration

- 4:30 Estimation of Sensitivity and Specificity of Multiple Tests and Disease Prevalence for Repeated Measures without Gold Standard**  
Chunling Wang\*, University of South Carolina; Timothy E. Hanson, Medtronic Inc.

- 4:45 A New Measure of Diagnostic Accuracy with Cut-point Selection Criterion for K-stage Diseases Using Concordance and Discordance**  
Jing Kersey\*, Hani Samawi, Jingjing Yin, Haresh Rochani, Xinyan Zhang and Chen Mo, Georgia Southern University

- 5:00 Floor Discussion**

## 58. CONTRIBUTED PAPERS: VARIABLE SELECTION

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Yet Nguyen, Old Dominion University

- 3:45 Variable Selection in Enriched Dirichlet Process with Applications to Causal Inference**  
Kumaresh Dhara\* and Michael J. Daniels, University of Florida

- 4:00 Model Confidence Bounds for Variable Selection**  
Yang Li\*, Renmin University of China

- 4:15 All Models are Wrong but Many are Useful: Variable Importance for Black-Box, Proprietary, or Unknown Prediction Models, with Model Class Reliance**  
Aaron J. Fisher\*, Harvard T. H. Chan School of Public Health; Cynthia Rudin, Duke University; Francesca Dominici, Harvard University

- 4:30 Variable Screening with Multiple Studies**  
Tianzhou Ma\*, University of Maryland, College Park; Zhao Ren and George Tseng, University of Pittsburgh

- 4:45 Simultaneous Estimation and Variable Selection for Interval-Censored Data with Broken Adaptive Ridge Regression**  
Qiwei Wu\*, University of Missouri, Columbia; Hui Zhao, Zhongnan University of Economics and Law, China; Gang Li, University of California, Los Angeles; Jianguo Sun, University of Missouri, Columbia

- 5:00 An Application of Penalized Quasi-Likelihood in Variable Selection on Parametric Accelerated Failure Time Models with Frailty**  
Sarbesh R. Pandeya\*, Lili Yu, Hani M. Samawi and Xinyan Zhang, Georgia Southern University

- 5:15 Simultaneous Selection and Inference for Varying Coefficients with Zero Regions: A Soft-Thresholding Approach**  
Yuan Yang\*, Jian Kang and Yi Li, University of Michigan

## 59. CONTRIBUTED PAPERS: CAUSAL INFERENCE

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Tianchen Qian, Harvard University

- 3:45 Causal Isotonic Regression**  
Ted Westling\*, University of Pennsylvania; Peter Gilbert, Fred Hutchinson Cancer Research Center; Marco Carone, University of Washington

- 4:00 Multiply Robust Two-Sample Instrumental Variable Estimation**  
BaoLuo Sun\*, National University of Singapore

- 4:15 An Instrumental Variable for Cox Models Extended to Non-Proportional Hazards and Effect Modification**  
James O'Malley\*, Pablo Martinez-Camblor and Todd MacKenzie, Geisel School of Medicine at Dartmouth; Douglas O. Staiger, Dartmouth College; Philip P. Goodney, Geisel School of Medicine at Dartmouth

- 4:30 Post-Randomization Biomarker Effect Modification in an HIV Vaccine Clinical Trial**  
Bryan S. Blette\*, University of North Carolina, Chapel Hill; Peter B. Gilbert, University of Washington; Bryan E. Shepherd, Vanderbilt University; Michael G. Hudgens, University of North Carolina, Chapel Hill

Denotes student award winner



# SCIENTIFIC PROGRAM

MONDAY, MARCH 25

## 4:45 Principal Stratification for Causal Effects Conditioning on a Binary Post-Treatment Variable in Clinical Trials

Judah Abberbock\*, GlaxoSmithKline; Gong Tang, University of Pittsburgh

## 5:00 Floor Discussion

## 60. CONTRIBUTED PAPERS: GENETIC EFFECTS/HERITABILITY

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jingwen Zhang, Harvard T.H. Chan School of Public Health

## 3:45 Phenotype Imputation Integrating GWAS Summary Association Statistics, Deep Phenotyped Cohorts and Large Biobanks: Application to Nicotine Uptake Phenotypes

Lina Yang\* and Dajiang Liu, Pennsylvania State College of Medicine

## 4:00 A Unified Method for Rare Variant Analysis of Gene-Environment Interactions

Elise Lim\*, Boston University; Han Chen, University of Texas Health Science Center at Houston; José Dupuis and Ching-Ti Liu, Boston University

## 4:15 Heritability Estimation and Genetic Association Testing in Longitudinal Twin Studies

Souvik Seal\* and Saonli Basu, University of Minnesota

## 4:30 Comparison of Hypothesis Testing Methods on Random Genetic Effects in Family Data

Nicholas DeVogel\* and Tao Wang, Medical College of Wisconsin

## 4:45 Small and Large Sample Bias of REML Estimates of Genomic Heritability Estimates: An Assessment using Big Data

Raka Mandal\*, Tapabrata Maiti and Gustavo De Los Campos, Michigan State University

## 5:00 An Adaptive Test for High-Dimensional Generalized Linear Models with Application to Detect Gene-Environment Interactions

Chong Wu\*, Florida State University; Gongjun Xu, University of Michigan; Xiaotong Shen and Wei Pan, University of Minnesota

## 5:15 Floor Discussion

## 61. CONTRIBUTED PAPERS: COMPUTATIONAL METHODS AND MASSIVE DATA SETS

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Zhipeng Wang, Genentech

## 3:45 On the Use of Optimal Transportation Theory to Merge Databases: Application to Clinical Research

Nicolas J. Savy\*, Toulouse Institute of Mathematics; Valérie Garès, INSA of Rennes; Chloé Dimeglio, CHU of Toulouse; Gregory Guernec and Benoit Lepage, INSERM unit 1027; Michael R. Kosorok, University of North Carolina, Chapel Hill; Philippe Saint-Pierre, Toulouse Institute of Mathematics

## 4:00 An Online Updating Approach for Testing the Proportional Hazards Assumption with Streams of Big Survival Data

Yishu Xue\*, HaiYing Wang, Jun Yan and Elizabeth D. Schifano, University of Connecticut

## 4:15 Real-Time Regression Analysis of Streaming Health Datasets

Lan Luo\* and Peter X.K. Song, University of Michigan

## 4:30 Application of Deep Convolutional Neural Networks in Classification of Protein Subcellular Localization with Microscopy Images

Mengli Xiao\*, Xiaotong Shen and Wei Pan, University of Minnesota

## 4:45 Extracting the Common Pattern between High-Dimensional Datasets

Zhe Qu\* and Mac Hyman, Tulane University

## 5:00 Projection Inference for Penalized Regression Estimators

Biyue Dai\* and Patrick Breheny, University of Iowa

## 5:15 Floor Discussion

Denotes student award winner

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

**TUESDAY, MARCH 26**

**8:30–10:15 A.M.**

## 62. RECENT ADVANCES IN BAYESIAN NETWORK META-ANALYSIS

*Franklin Hall 1, 4th Floor*

**SPONSOR:** ENAR, ASA Bayesian Statistical Science Section, ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Health Policy Statistics Section

**ORGANIZER:** Haitao Chu, University of Minnesota

**CHAIR:** Jincheng Zhou, University of Minnesota

### 8:30 Bayesian Network Meta-Regression for Ordinal Outcomes: Applications to Comparing Crohn's Disease Treatments

Joseph G. Ibrahim\*, University of North Carolina, Chapel Hill; Yeongjin Gwon and Ming-Hui Chen, University of Connecticut; May Mo, Tony Jiang and Amy Xia, Amgen Inc.

### 8:55 Bayesian Joint Network Meta-Regression Methods Adjusting for Post-Randomization Variables

Jing Zhang\* and Mark Wymer, University of Maryland; Qinshu Lian, Genentech; Haitao Chu, University of Minnesota

### 9:20 Bayesian Inconsistency Detection for Network Meta-Analysis

Ming-Hui Chen\*, Hao Li and Cheng Zhang, University of Connecticut; Joseph G. Ibrahim, University of North Carolina, Chapel Hill; Arvind K. Shah and Jianxin Lin, Merck & Co.

### 9:45 Bayesian Hierarchical Models for N-of-1 Trials with Ordinal Outcomes

Youdan Wang\* and Christopher Schmid, Brown University

### 10:10 Floor Discussion

## 63. STATISTICAL METHODS TO SUPPORT VALID AND EFFICIENT USE OF ELECTRONIC HEALTH RECORDS DATA

*Franklin Hall 2, 4th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistical Learning and Data Science Section

**ORGANIZER:** Guanhua Chen, University of Wisconsin, Madison

**CHAIR:** Guanhua Chen, University of Wisconsin, Madison

### 8:30 Enabling Imprecise EHR Data for Precision Medicine

Tianxi Cai\*, Harvard University

### 8:55 A Bayesian Nonparametrics for Missing Data in EHRs

Michael J. Daniels\*, University of Florida

### 9:20 Sampling Designs for Resource Efficient Collection of Outcome Labels for Statistical Learning, with Applications to Electronic Medical Records

Patrick J. Heagerty\* and Wei Ling Katherine Tan, University of Washington

### 9:45 Accounting for Differential Misclassification in EHR-Derived Phenotypes

Rebecca A. Hubbard\*, University of Pennsylvania

### 10:10 Floor Discussion

## 64. RECENT ADVANCES IN THE ANALYSIS OF TIME-TO-EVENT OUTCOMES SUBJECT TO A TERMINAL EVENT

*Franklin Hall 13, 4th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Epidemiology Section

**ORGANIZER:** Sebastien Haneuse, Harvard T.H. Chan School of Public Health

**CHAIR:** Sebastien Haneuse, Harvard T.H. Chan School of Public Health

### 8:30 Flexible Accelerated Time Modeling of Recurrent Events Data in the Presence of a Dependent Terminal Event

Limin Peng\* and Bo Wei, Emory University; Zhumin Zhang and HuiChuan Lai, University of Wisconsin, Madison

### 8:55 Bayesian Variable Selection for a Semi-Competing Risks Model with Three Hazard Functions

Andrew G. Chapple\*, Louisiana State University School of Public Health; Marina Vannucci, Rice University; Peter F. Thall and Steven Lin, University of Texas MD Anderson Cancer Center

### 9:20 A Generalized Nested Case-Control Design for the Semi-Competing Risks Setting

Ina Jazic\*, Vertex Pharmaceuticals; Tianxi Cai and Sebastien Haneuse, Harvard T.H. Chan School of Public Health

### 9:45 Analysis of Semi-Competing Risks Data via Bivariate Longitudinal Models

Daniel Nevo\*, Tel Aviv University; Sebastien Haneuse, Harvard T.H. Chan School of Public Health

### 10:10 Floor Discussion

## 65. RECENT ADVANCES IN STATISTICAL METHODS FOR PRECISION MEDICINE

*Salon B, 5th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section

**ORGANIZERS:** Jian Kang and Kevin He, University of Michigan

**CHAIR:** Kevin He, University of Michigan

### 8:30 Variable Selection in Joint Frailty Models of Recurrent and Terminal Events

Lei Liu\*, Washington University in St. Louis; Dongxiao Han and Liuquan Sun, Chinese Academy of Sciences; Xiaogang Su, University of Texas at El Paso; and Zhou Zhang, Northwestern University Feinberg School of Medicine

### 8:55 Subgroup Identification Using Electronic Health Record Data

Marianthi Markatou\*, State University of New York at Buffalo

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

**9:20 Human Disease Network (HDN) Analysis of Disease Prevalence**  
Shuangge Ma\*, Yale University

**9:45 Modeling Time-Varying Effects of Multilevel Risk Factors of Hospitalizations in Patients on Dialysis**  
Damla Senturk\* and Yihao Li, University of California, Los Angeles; Danh V. Nguyen, Yanjun Chen, Connie M. Rhee and Kamyar Kalantar-Zadeh, University of California, Irvine

**10:10 Floor Discussion**

## 66. CHALLENGES AND ADVANCES IN WEARABLE TECHNOLOGY

*Salon C, 5th Floor*

**SPONSOR:** ENAR, ASA Health Policy Statistics Section

**ORGANIZER:** Ekaterina Smirnova, Virginia Commonwealth University

**CHAIR:** Elizabeth J. McGuffey, Indiana University, Bloomington

**8:30 Classification of Human Activity Based on the Raw Accelerometry Data: Comparison of the Data Transformations**  
Jaroslaw Harezlak\* and Marcin Straczekiewicz, Indiana University, Bloomington; Jacek Urbanek, Johns Hopkins University

**8:55 Sub-Second Level Accelerometry Data in Health Research: Challenges and Opportunities**  
Jacek K. Urbanek\*, Johns Hopkins University School of Medicine; Marta Karas, Johns Hopkins Bloomberg School of Public Health; Marcin Straczekiewicz and Jaroslaw Harezlak, Indiana University, Bloomington; Jiawei Bai, Vadim Zipunnikov and Ciprian Crainiceanu, Johns Hopkins Bloomberg School of Public Health

**9:20 Physical Activity versus Inactivity versus Sleep: Nonparametric Estimates of Isotemporal Substitution Effects**  
John W. Staudenmayer\*, University of Massachusetts Amherst

**9:45 Deriving Objective Activity Measures from Accelerometry Data for Mortality Prediction Models**  
Ekaterina Smirnova\*, Virginia Commonwealth University; Andrew Leroux and Ciprian M. Crainiceanu, Johns Hopkins Bloomberg School of Public Health

**10:10 Floor Discussion**

## 67. STATISTICAL MODELING IN CELL BIOLOGY

*Salon D, 5th Floor*

**SPONSOR:** IMS

**ORGANIZER:** Vladimir Minin, University of California, Irvine

**CHAIR:** Jessica Li, University of California, Los Angeles

**8:30 Missing Data and Technical Variability in Single-Cell RNA-Sequencing Experiments**  
Stephanie Hicks\*, Johns Hopkins Bloomberg School of Public Health; William Townes, Harvard T.H. Chan School of Public Health; Martin Aryee, Massachusetts General Hospital; Rafael Irizarry, Dana-Farber Cancer Institute

**8:55 Fitting Stochastic Models to in Vivo Cell Lineage Tracking Data**  
Jason Xu\*, Duke University; Samson Koelle and Peter Gutter, University of Washington; Chuanfeng Wu and Cynthia Dunbar, National Heart, Lung, and Blood Institute, National Institutes of Health; Janis Abkowitz, University of Washington; Vladimir Minin, University of California, Irvine

**9:20 Imputation of Single-Cell Gene Expression with Autoencoder Neural Networks**  
Audrey Q. Fu\*, University of Idaho

**9:45 Models for Dependent Data in Single Cell Assays**  
Andrew McDavid\* and Corey Kimzey, University of Rochester

**10:10 Floor Discussion**

## 68. CONTRIBUTED PAPERS: DIAGNOSTICS, ROC, AND RISK PREDICTION

*405, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Gene Anthony Pennello, U.S. Food and Drug Administration

**8:30 On Kullback-Leibler Divergence as a Measure for Medical Diagnostics and Cut-Point Selection Criterion**  
Hani Samawi\*, Jingjing Yin, Xinyan Zhang, Lili Yu, Hareesh Rochani and Robert Vogel, Georgia Southern University

**8:45 Receiver Operating Characteristic Curves and Confidence Bands for Support Vector Machines**  
Daniel J. Lockett\*, University of North Carolina, Chapel Hill; Eric B. Laber, North Carolina State University; Michael R. Kosorok, University of North Carolina, Chapel Hill

**9:00 Proper ROC Models: Dual Beta Model and Weighted Power Function Model**  
Hongying Peng\*, Douglas Mossman and Marepalli Rao, University of Cincinnati

**9:15 A General Framework for using the Overall Concordance Statistic to Assess the Discriminatory Ability of Risk Predictions**  
Li C. Cheung\*, National Cancer Institute, National Institutes of Health; Qing Pan, The George Washington University; Barry Graubard, National Cancer Institute, National Institutes of Health



# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 9:30 Evaluating Predictive Accuracy of Traditional and Machine Learning Based Survival Models

Yue-Ming Chen\* and Dai Feng, Merck Research Laboratories; Nicholas C. Henderson, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University; Vladimir Svetnik, Merck Research Laboratories

## 9:45 Floor Discussion

## 69. CONTRIBUTED PAPERS: MICROBIOME DATA: FINDING ASSOCIATIONS AND TESTING

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Shuang Jiang, Southern Methodist University

## 8:30 Robust Screening for Associations Between Microbiome Community Profiles and a Large Number of Individual Genomic Outcomes

Weijia Fu\*, Nanxun Ma and Michael C. Wu, Fred Hutchinson Cancer Research Center

## 8:45 Improved Variance Component Score Tests of Marker-Environment Interactions

Nanxun Ma\*, University of Washington; Michael C. Wu and Jing Ma, Fred Hutchinson Cancer Research Center

## 9:00 Ecological Dissimilarities for Paired and Longitudinal Microbiome Association Analysis

Anna M. Plantinga\*, Williams College; Michael C. Wu, Fred Hutchinson Cancer Research Center

## 9:15 An Adaptive Distance-Based Kernel Association Test Based on the Generalized Linear Mixed Effect Model for Correlated Microbiome Studies

Hyunwook Koh\* and Ni Zhao, Johns Hopkins Bloomberg School of Public Health

## 9:30 An Adaptive Multivariate Two-Sample Test with Application to Microbiome Differential Abundance Analysis

Xiang Zhan\*, The Pennsylvania State University

## 9:45 Semiparametric Methods for Testing for Differential Abundance in Microbiome Studies

Olivier Thas\*, Hasselt University, Belgium, Ghent University, Belgium and University of Wollongong, Australia; Leyla Kodalcı, Hasselt University, Belgium; Stijn Hawinkel, Ghent University, Belgium

## 10:00 Floor Discussion

## 70. CONTRIBUTED PAPERS: COMPARATIVE EFFECTIVENESS, CLUSTERED AND CATEGORICAL DATA

Franklin Hall 3, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Lili Zhao, University of Michigan

## 8:30 A Potential Outcomes Approach to Subgroup Discovery in Cost-Effectiveness Analysis

Nicholas A. Illenberger\* and Nandita Mitra, University of Pennsylvania; Andrew J. Spieker, Vanderbilt University Medical Center

## 8:45 Estimating Treatment Importance in Multidrug-Resistant Tuberculosis Using Targeted Learning: An Observational Individual Patient Data Meta-Analysis

Guanbo Wang\*, McGill University

## 9:00 Limited Information Empirical Bayes for Classification of Subjects Across Conditions

Hillary Koch\*, The Pennsylvania State University; Siqi Xiang, University of North Carolina, Chapel Hill; Han Wang, Zhejiang University; Feipeng Zhang, Hunan Normal University; Qunhua Li, The Pennsylvania State University

## 9:15 Conditional Log-linear Regression for Stratified Pairs of Ordinal Responses

Jing Yu\* and Gary Koch, University of North Carolina, Chapel Hill

## 9:30 Latent Class Model for Finding Co-Occurrent Patterns in Process Data

Guanhua Fang\*, Zhiliang Ying and Jingchen Liu, Columbia University

## 9:45 Accuracy of Latent Class Item Response Theory Models for Examining Measurement Invariance in Patient-Reported Outcomes Measures

Tolulope Sajobi\*, University of Calgary; Richard Sawatzky, Trinity Western University; Lara Russell, University of British Columbia; Oluwaseyi A. Lawal, University of Calgary; Juxin Liu, University of Saskatchewan; Bruno D. Zumbo, University of British Columbia; Lisa M. Lix, University of Manitoba

## 10:00 Floor Discussion

## 71. CONTRIBUTED PAPERS: CAUSAL INFERENCE AND MEASUREMENT ERROR

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Hwanhee Hong, Duke University

## 8:30 Instrumental Variable Approach to Estimating the Scalar-on-Function Regression Model with Measurement Error with Application to Energy Expenditure Assessment in Childhood Obesity

Carmen D. Tekwe\* and Roger S. Zoh, Texas A&M University; Lan Xue, Oregon State University

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 8:45 Calibrating Validation Samples when Correcting for Measurement Error in Intervention Study Outcomes

Benjamin Ackerman\*, Johns Hopkins Bloomberg School of Public Health;  
Juned Siddique, Northwestern University Feinberg School of Medicine;  
Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health

## 9:00 Assessing the Impact of Differential Measurement Error on Outcomes in a Longitudinal Lifestyle Intervention Study

David A. Aaby\* and Juned Siddique, Northwestern University Feinberg School of Medicine

## 9:15 Comparison of Causal Methods for Average Treatment Effect Estimation Allowing Covariate Measurement Error

Zhou Feng\* and Yi Huang, University of Maryland, Baltimore County

## 9:30 Estimation of Natural Indirect Effects Robust to Unmeasured Confounding and Mediator Measurement Error

Isabel R. Fulcher\* and Xu Shi, Harvard University; Eric J. Tchetgen Tchetgen, The Wharton School, University of Pennsylvania

## 9:45 Floor Discussion

## 72. CONTRIBUTED PAPERS: GENOMICS, PROTEOMICS, OR OTHER OMICS

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Zhou Fang, Brigham and Women's Hospital, Harvard Medical School

## 8:30 A Novel Test for Positive Selection using Protein Structural Information

Peter B. Chi\*, Villanova University; David A. Liberles, Temple University

## 8:45 Penalized Multiple Co-Inertia Analysis with Application to Integrative Analysis of Multi -Omics Data

Eun Jeong Min\* and Qi Long, University of Pennsylvania

## 9:00 Understanding and Predicting Rapid Disease Progression in the Presence of Sparse Effects: A Case Study with Cystic Fibrosis Lung Function and Proteomics Data

Emrah Gecili\*, Cincinnati Children's Hospital and University of Cincinnati;  
John P. Clancy, Cystic Fibrosis Foundation; Rhonda Szczesniak and Assem Ziady, Cincinnati Children's Hospital and University of Cincinnati

## 9:15 MOVIE: Multi-Omics Visualization of Estimated Contributions

Sean D. McCabe\*, Dan-Yu Lin and Michael I. Love, University of North Carolina, Chapel Hill

## 9:30 Improved Detection of Epigenetic Marks with Mixed Effects Hidden Markov Models

Pedro L. Baldoni\*, Naim U. Rashid and Joseph G. Ibrahim, University of North Carolina, Chapel Hill

## 9:45 Accurate and Efficient Estimation of Small P-values with the Cross-Entropy Method: Applications in Genomic Data Analysis

Yang Shi\*, Augusta University; Hui Jiang, University of Michigan; Huining Kang, University of New Mexico Comprehensive Cancer Center; Ji-Hyun Lee, University of Florida

## 10:00 Second-Generation p-values in a High Dimensional Analysis of Prostate Cancer Variants

Valerie F. Welty\*, Robert A. Greevy, Jeffrey R. Smith, William D. Dupont and Jeffrey D. Blume, Vanderbilt University

## 73. CONTRIBUTED PAPERS: IMAGING METHODS

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Joanne C. Beer, University of Pennsylvania

## 8:30 Hierarchical Shrinkage Priors for Using Images as Predictors in Bayesian Generalized Linear Models

Justin M. Leach\*, Inmaculada Aban and Nengjun Yi, University of Alabama at Birmingham

## 8:45 A Simplified Crossing Fiber Model in Diffusion Weighted Imaging

Kaushik Ghosh\*, University of Nevada Las Vegas; Sheng Yang, Case Western Reserve University; Ken Sakaie, Cleveland Clinic; Satya S. Sahoo, Case Western Reserve University; Sarah J. Carr, King's College, London; Curtis Tatsuoka, Case Western Reserve University

## 9:00 Spectral Analysis of Brain Signals: A New Bayesian Nonparametric Approach

Guillermo Cuauhtemoc Granados-Garcia\*, King Abdullah University of Science and Technology; Mark Fiecas, University of Minnesota; Hernando Ombao, King Abdullah University of Science and Technology

## 9:15 Copula Modeling of Spectral Decompositions of Multivariate Non-Stationary Time Series

Yongxin Zhu\* and Charles Fontaine, King Abdullah University of Science and Technology; Ron Frostig, University of California, Irvine; Hernando Ombao, King Abdullah University of Science and Technology

## 9:30 Comparing Summary Methods and a Spatiotemporal Model in the Analysis of Longitudinal Imaging Data

Brandon J. George\*, Thomas Jefferson University; Inmaculada B. Aban, University of Alabama at Birmingham

## 9:45 A Spatial Bayesian Semi-Parametric Model of Positive Definite Matrices for Diffusion Tensor Imaging

Zhou Lan\* and Brian J. Reich, North Carolina State University; Dipankar Bandyopadhyay, Virginia Commonwealth University

## 10:00 Floor Discussion

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

**TUESDAY, MARCH 26**

**10:15—10:30 A.M.**

## REFRESHMENT BREAK WITH OUR EXHIBITORS

*Franklin Hall Foyer, 4th Floor*

**TUESDAY, MARCH 26**

**10:30 A.M.—12:15 P.M.**

### 74. PRESIDENTIAL INVITED ADDRESS

*Salons E-F, 5th Floor*

**SPONSOR:** ENAR

**ORGANIZER/CHAIR:** Sarah J. Ratcliffe, University of Virginia

**10:30** Introduction

**10:35** Distinguished Student Paper Awards

**10:45** **A Particulate Solution: Data Science in the Fight to Stop Air Pollution and Climate Change**

Francesca Dominici, Ph.D, Clarence James Gamble Professor of Biostatistics, Population and Data Science, Harvard T.H. Chan School of Public Health and Co-Director, Data Science Initiative, Harvard University

**TUESDAY, MARCH 26**

**1:45—3:30 P.M.**

### 75. RESOURCE EFFICIENT STUDY DESIGNS FOR OBSERVATIONAL AND CORRELATED DATA

*Franklin Hall 1, 4th Floor*

**SPONSOR:** ENAR, IMS, ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Jonathan Schildcrout, Vanderbilt University Medical Center

**CHAIR:** Sebastien Haneuse, Harvard T.H. Chan School of Public Health

**1:45** **Multi-Wave, Response-Selective Study Designs for Longitudinal Binary Data**

Nathaniel D. Mercaldo\*, Massachusetts General Hospital and Harvard University; Jonathan S. Schildcrout, Vanderbilt University Medical Center

**2:10** **Efficient Design and Analysis of Two-Phase Studies with a Longitudinal Continuous Outcome**

Ran Tao\*, Vanderbilt University Medical Center

**2:35** **Response Driven Study Designs for Multivariate Longitudinal Data**

Jonathan S. Schildcrout\*, Vanderbilt University Medical Center

**3:00** **Semiparametric Generalized Linear Models: Application to Biased Samples**

Paul Rathouz\*, Dell Medical School, University of Texas, Austin

**3:25** Floor Discussion

### 76. RECENT ADVANCES IN THE STUDY OF INTERACTION

*Franklin Hall 2, 4th Floor*

**SPONSOR:** ENAR

**ORGANIZERS:** Eric Tchetgen Tchetgen, University of Pennsylvania and Xu Shi, Harvard University

**CHAIR:** Eric Tchetgen Tchetgen, University of Pennsylvania

**1:45** **The Interaction Continuum**

Tyler J. VanderWeele\*, Harvard University

**2:15** **Additive versus Multiplicative Interaction: The Epidemiological Folklore regarding Heterogeneity across Studies**

Bhramar Mukherjee\*, University of Michigan

**2:45** **A General Approach to Detect Gene (G)-environment (E) Additive Interaction Leveraging G-E Independence in Case-Control Studies**

Xu Shi\*, Harvard University; Eric J. Tchetgen Tchetgen, Wharton School of the University of Pennsylvania; Tamar Sofer, Brigham and Women's Hospital and Harvard Medical School; Benedict H. W. Wong, Harvard University

**3:15** **Discussant:**

James Dai, Fred Hutchinson Cancer Research Center

**3:25** Floor Discussion

### 77. EXPANDING RANK TESTS: ESTIMATES, CONFIDENCE INTERVALS, MODELING, AND APPLICATIONS

*Salon B, 5th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section

**ORGANIZER:** Michael P. Fay, National Institute of Allergy and Infectious Diseases, National Institutes of Health

**CHAIR:** Sally Hunsberger, National Institute of Allergy and Infectious Diseases, National Institutes of Health

**1:45** **Confidence Intervals and Causal Inference with the Mann-Whitney Parameter that are Compatible with the Wilcoxon-Mann-Whitney Test**

Michael P. Fay\*, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Yaakov Malinovsky, University of Maryland, Baltimore County; Erica H. Brittain, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Joanna H. Shih, National Cancer Institute, National Institutes of Health; Dean A. Follmann, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Erin E. Gabriel, Karolinska Institute, Stockholm, Sweden



# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 2:10 Small Sample Inference for Probabilistic Index Models

Jan De Neve\* and Gustavo Amorim, Ghent University; Olivier Thas, Ghent University, Hasselt University and University of Wollongong; Karel Vermeulen, Ghent University; Stijn Vansteelandt, Ghent University and London School of Hygiene and Tropical Medicine

## 2:35 Rank-Based Procedures in Factorial Designs: Hypotheses about Nonparametric Treatment Effects

Frank Konietzschke\*, University of Texas, Dallas

## 3:00 Desirability of Outcome Ranking (DOOR): Motivation and Examples

Scott R. Evans\*, The George Washington University

## 3:25 Floor Discussion

## 78. NOVEL STATISTICAL METHODS TO ANALYZE SELF-REPORTED OUTCOMES SUBJECT TO RECALL ERROR IN OBSERVATIONAL STUDIES

Franklin Hall 13, 4th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Epidemiology Section

**ORGANIZER:** Sedigheh Mirzaei Salehabadi, St. Jude Children's Research Hospital

**CHAIR:** Sedigheh Mirzaei Salehabadi, St. Jude Children's Research Hospital

## 1:45 Analyzing Recurrent Events in Presence of Recall Error: An Application to Time-to-Hospitalization

Rajeshwari Sundaram\*, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health; Sedigheh Mirzaei, St. Jude Children's Research Hospital; Edwina Yeung, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health

## 2:10 Estimation of Menarcheal Age Distribution from Imperfectly Recalled Data

Debasis Sengupta\*, Indian Statistical Institute, Kolkata; Sedigheh Mirzaei Salehabadi, St. Jude Children's Research Hospital; Rahul Ghosal, North Carolina State University

## 2:35 Variable Selection in High Dimensional Datasets in the Presence of Self-Reported Outcomes

Raji Balasubramanian\*, University of Massachusetts Amherst; Mahlet G. Tadesse, Georgetown University; Andrea S. Foulkes, Mount Holyoke College; Yunsheng Ma, University of Massachusetts Medical School; Xiangdong Gu, University of Massachusetts Amherst

## 3:00 Measurement Error Correction in Longitudinal Dietary Intervention Studies in the Presence of Nonignorable Missing Data

Juned Siddique\*, Caroline P. Groth and David A. Aaby, Northwestern University; Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health; Michael J. Daniels, University of Florida

## 3:25 Floor Discussion

## 79. STATISTICAL ADVANCE IN HUMAN MICROBIOME DATA ANALYSIS

Salon C, 5th Floor

**SPONSOR:** ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Huilin Li, New York University

**CHAIR:** Huilin Li, New York University

## 1:45 Testing Hypotheses about Associations between Taxonomic Groupings and Traits using 16S rRNA Data: A Bottom-up Approach.

Glen A. Satten\*, Centers for Disease Control and Prevention; Yijuan Hu and Yunxiao Li, Emory University

## 2:10 Beta-Diversity Discriminatory Power: Comparison of PERMANOVA, Mirkat, and Using Standard Microbiome Reference Groups

Mitchell Henry Gail\* and Yunhu Wan, National Cancer Institute, National Institutes of Health

## 2:35 Zero-Inflated Logistic Normal Model for Analyzing Microbiome Relative Abundance Data

Zhigang Li\*, University of Florida

## 3:00 Interactive Statistical Analysis of Host-Microbiome Interaction

Hector Corrada Bravo\*, University of Maryland, College Park

## 3:25 Floor Discussion

## 80. STATISTICAL MEDIATION ANALYSIS FOR HIGH-DIMENSIONAL DATA

Salon D, 5th Floor

**SPONSOR:** IMS

**ORGANIZER:** Alexander Alekseyenko, Medical University of South Carolina

**CHAIR:** Alexander Alekseyenko, Medical University of South Carolina

## 1:45 Learning Causal Networks via Additive Faithfulness

Kuang-Yao Lee\*, Temple University; Tianqi Liu, Google LLC; Bing Li, The Pennsylvania State University; Hongyu Zhao, Yale University

## 2:10 Statistical Methodology for High Dimensional Mediation Analysis

Andriy Derkach\* and Joshua Sampson, National Cancer Institute, National Institutes of Health

## 2:35 Hypothesis Testing in High-Dimensional Instrumental Variables Regression with an Application to Genomics Data

Jiarui Lu\* and Hongzhe Li, University of Pennsylvania

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 3:00 Mediation Analysis in Genomic Settings in the Presence of Reverse Causality and other Challenges

Joshua Millstein\*, University of Southern California

## 3:25 Floor Discussion

### 81. CONTRIBUTED PAPERS: PREDICTION AND PROGNOSTIC MODELING

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Zhe Chen, Novartis Pharmaceuticals Corporation

## 1:45 Dynamic Risk Prediction Using Longitudinal Biomarkers of High Dimensions

Lili Zhao\* and Susan Murray, University of Michigan

## 2:00 Estimating Disease Onset from Change Points of Markers Measured with Error

Unkyung Lee\* and Raymond J. Carroll, Texas A&M University; Karen Marder, Columbia University Medical Center; Yuanjia Wang, Columbia University; Tanya P. Garcia, Texas A&M University

## 2:15 Development of an International Prostate Cancer Risk Tool Integrating Data from Multiple Heterogeneous Cohorts

Donna Pauler Ankerst\* and Johanna Straubinger, Technical University Munich

## 2:30 Predicting Service Use and Functioning for People with First Episode Psychosis in Coordinated Specialty Care

Melanie M. Wall\*, Jenn Scodes and Cale Basaraba, Columbia University

## 2:45 Race and Gender Specific Statistical Comparison of Growth Curve Models in First Years of Life

Mehmet Kocak\* The University of Tennessee Health Science Center; Alemayehu Wolde, University of Memphis; Frances A. Tylavsky, The University of Tennessee Health Science Center

## 3:00 Floor Discussion

### 82. CONTRIBUTED PAPERS: ADAPTIVE DESIGNS FOR CLINICAL TRIALS

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Yeonhee Park, Medical University of South Carolina

## 1:45 Validity and Robustness of Tests in Survival Analysis under Covariate-Adaptive Randomization

Ting Ye\* and Jun Shao, University of Wisconsin, Madison

## 2:00 Platform Trial Designs for Borrowing Adaptively from Historical Control Data

James Paul Normington\*, University of Minnesota

## 2:15 Improved Estimates for RECIST Responder Rates in the Small Treatment Arms in Platform Screening Trials

James Dunyak\* and Nidal Al-Huniti, AstraZeneca

## 2:30 A Continuous Reassessment Method for Pediatric Phase I Clinical Trials

Yimei Li\*, University of Pennsylvania; Ying Yuan, University of Texas MD Anderson Cancer Center

## 2:45 A Simulation-Based Sample Size Determination for Adaptive Seamless Phase II/III Design

Zhongying Xu\*, John A. Kellum, Gary M. Marsh and Chung-Chou H. Chang, University of Pittsburgh

## 3:00 A Bayesian Adaptive Basket Trial Design for Related Diseases using Heterogeneous Endpoints

Matthew Austin Psioda\*, Joseph G. Ibrahim and Jiawei Xu, University of North Carolina, Chapel Hill; Tony Jiang and Amy Xia, Amgen

## 3:15 Adaptively Monitoring Clinical Trials with Second-Generation p-values

Jonathan J. Chipman\*, Robert A. Greevy, Jr., Lindsay Mayberry and Jeffrey D. Blume, Vanderbilt University

### 83. CONTRIBUTED PAPERS: BAYESIAN APPROACHES TO SURVEYS AND SPATIO-TEMPORAL MODELING

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Jasmit S. Shah, Aga Khan University

## 1:45 Estimating of Prostate Cancer Incidence Rates Using Serially Correlated Generalized Multivariate Models

Manoj Pathak\*, Murray State University; Jane L. Meza, University of Nebraska Medical Center; Kent M. Eskridge, University of Nebraska, Lincoln

## 2:00 Bayesian Hierarchical Models for Voxel-Wise Classification of Prostate Cancer Accounting for Spatial Correlation and Between-Patient Heterogeneity in the Multi-Parametric MRI Data

Jin Jin\*, Lin Zhang, Ethan Leng, Gregory J. Metzger and Joseph S. Koopmeiners, University of Minnesota

## 2:15 A Bayesian Shape Invariant Growth Curve Model for Longitudinal Data

Mohammad Alfrad Nobel Bhuiyan\*, Heidi Sucharew and Md Monir Hossain, Cincinnati Children's Hospital Medical Center

## 2:30 Bayesian Record Linkage Under Limited Linking Information

Mingyang Shan\*, Kali Thomas and Roe Gutman, Brown University

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 2:45 Bayesian Variable Selection in Growth Mixture Model with Missing Covariates Data

Zihang Lu\* and Wendy Lou, University of Toronto

## 3:00 Floor Discussion

## 84. CONTRIBUTED PAPERS: CAUSAL EFFECTS WITH PROPENSITY SCORES/WEIGHTING/MATCHING

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Amanda H. Pendegraft, University of Alabama at Birmingham

## 1:45 Using Propensity Scores with Treatment Selection Biomarkers

Hulya Kocyigit\*, University of Georgia

## 2:00 Building Representative Matched Samples with Multi-Valued Treatments

Magdalena Bennett\*, Columbia University; Juan Pablo Vielma, Massachusetts Institute of Technology; Jose R. Zubizarreta, Harvard University

## 2:15 Triplet Matching for Estimating Causal Effects with Three Treatment Arms: A Comparative Study of Mortality by Trauma Center Level

Giovanni Nattino\* and Bo Lu, The Ohio State University; Junxin Shi, The Research Institute of Nationwide Children's Hospital; Stanley Lemeshow, The Ohio State University; Henry Xiang, The Research Institute of Nationwide Children's Hospital

## 2:30 A Novel Propensity Score Framework for a Continuous Treatment using the Cumulative Distribution Function

Derek W. Brown\*, University of Texas Health Science Center at Houston; Thomas J. Greene, GlaxoSmithKline; Michael D. Swartz, Anna V. Wilkinson and Stacia M. DeSantis, University of Texas Health Science Center at Houston

## 2:45 Minimal Dispersion Approximately Balancing Weights: Asymptotic Properties and Practical Considerations

Jose Zubizarreta\*, Harvard University; Yixin Wang, Columbia University

## 3:00 Conducting Mendelian Randomization Analysis on Summary Data under Case-Control Studies

Han Zhang\* and Lu Deng, National Cancer Institute, National Institutes of Health; Jing Qin, National Institute of Allergy and Infectious Diseases, National Institutes of Health; Kai Yu, National Cancer Institute, National Institutes of Health

## 85. CONTRIBUTED PAPERS: META-ANALYSIS

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Tianzhou Ma, University of Maryland, College Park

## 1:45 Bayesian Network Meta-Analysis of Treatment Toxicities

Aniko Szabo\* and Binod Dhakal, Medical College of Wisconsin

## 2:00 Adaptive Weighting Methods for Identifying Concordant Differentially Expressed Genes in Omics Meta-Analysis

Chien-Wei Lin\*, Medical College of Wisconsin; George C. Tseng, University of Pittsburgh

## 2:15 A Bayesian Hierarchical Model Estimating CACE in Meta-Analysis of Randomized Clinical Trials with Noncompliance

Jincheng Zhou\*, Haitao Chu, James S. Hodges and M. Fareed K. Suri, University of Minnesota

## 2:30 Quantifying the Evidence of Selective Publishing in Network Meta-Analysis: An EM Algorithm-Based Approach

Arielle K. Marks-Anglin\*, University of Pennsylvania; Jin Piao, University of Southern California; Jing Ning, University of Texas MD Anderson Cancer Center; Chongliang Luo and Yong Chen, University of Pennsylvania

## 2:45 High Resolution Fine-Mapping of 406 Smoking/Drinking Loci via a Novel Method that Synthesizes the Analysis of Exome-Wide and Genome-Wide Association Statistics

Yu Jiang\* and Daijiang Liu, Penn State College of Medicine

## 3:00 Bayesian Network Meta-Regression for Ordinal Outcomes Incorporating High-Dimensional Random Effects

Yeongjin Gwon\*, University of Nebraska Medical Center; Ming-Hui Chen, University of Connecticut; May Mo, Jiang Xun and Amy Xia, Amgen Inc.; Joseph Ibrahim, University of North Carolina, Chapel Hill

## 3:15 Multivariate Meta-Analysis of Randomized Controlled Trials with the Difference in Restricted Mean Survival Times

Isabelle R. Weir\*, Boston University School of Public Health; Lu Tian, Stanford University; Ludovic Trinquart, Boston University School of Public Health

## 86. CONTRIBUTED PAPERS: IMAGING APPLICATIONS AND TESTING

Franklin Hall 3, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Kaushik Ghosh, University of Nevada Las Vegas

Denotes student award winner

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

**1:45 Semiparametric Modeling of Time-Varying Activation and Connectivity in Task-Based fMRI Data**

Jun Young Park\*, University of Minnesota; Joerg Polzehl, Weierstrass Institute for Applied Analysis and Stochastics; Snigdhasu Chatterjee, University of Minnesota; André Brechmann, Leibniz-Institute for Neurobiology; Mark Fiecas, University of Minnesota

**2:00 On Predictability and Reproducibility of Individual Functional Connectivity Networks from Clinical Characteristics**

Emily L. Morris\* and Jian Kang, University of Michigan

**2:15 On Statistical Tests of Functional Connectome Fingerprinting**

Zeyi Wang\*, Johns Hopkins Bloomberg School of Public Health; Haris Sair, Johns Hopkins University School of Medicine; Ciprian Crainiceanu and Martin Lindquist, Johns Hopkins Bloomberg School of Public Health; Bennett A. Landman, Vanderbilt University; Susan Resnick, National Institute on Aging, National Institutes of Health; Joshua T. Vogelstein, Johns Hopkins University School of Medicine; Brian Caffo, Johns Hopkins Bloomberg School of Public Health

**2:30 A Comparison of Machine Learning Algorithms for Predicting Age from Multimodal Neuroimaging Data**

Joanne Beer\*, University of Pennsylvania; Helmet Karim, Dana Tudorascu, Howard Aizenstein, Stewart Anderson and Robert Krafty, University of Pittsburgh

**2:45 A Statistical Model for Stochastic Radiographic Lung Change Following Radiotherapy of Lung Cancer**

Nitai Das Mukhopadhyay\* and Viviana A. Rodriguez, Virginia Commonwealth University

**3:00 An Inter-Feature Correlation Guided Classifier for Alzheimer's Disease Prediction**

Yanming Li\*, University of Michigan

**3:15 Floor Discussion**

**TUESDAY, MARCH 26**

**3:30—3:45 P.M.**

## REFRESHMENT BREAK WITH OUR EXHIBITORS

*Franklin Hall Foyer, 4th Floor*

**TUESDAY, MARCH 26**

**3:45—5:30 P.M.**

### 87. METHODOLOGICAL CHALLENGES AND OPPORTUNITIES IN MENTAL HEALTH RESEARCH

*Franklin Hall 1, 4th Floor*

**SPONSOR:** ENAR, ASA Mental Health Statistics Section

**ORGANIZER:** Eva Petkova, New York University

**CHAIR:** Eva Petkova, New York University

**3:45 Integrative Learning to Combine Individualized Treatment Rules from Multiple Randomized Trials**

Yuanjia Wang\*, Columbia University

**4:10 Mixed-Effects Modeling to Compare Dynamic Treatment Regimens with SMART Data**

Brook Luers\*, University of Michigan; Min Qian, Columbia University; Inbal Nahum-Shani, University of Michigan; Connie Kasari, University of California, Los Angeles; Daniel Almirall, University of Michigan

**4:35 Handling Missing Clinical and Multimodal Imaging Data in Integrative Analysis with Applications to Mental Health Research**

Adam Ciarleglio\*, The George Washington University; Eva Petkova, New York University

**5:00 Design and Analytic Tools for Personalizing Healthcare Experiments**

Christopher H. Schmid\*, Brown University

**5:25 Floor Discussion**

### 88. NOVEL APPROACHES FOR GROUP TESTING FOR ESTIMATION IN BIOSTATISTICS

*Salon B, 5th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Section: Statistics and the Environment, ASA Statistics in Epidemiology Section

**ORGANIZER:** Paul Albert, National Cancer Institute, National Institutes of Health

**CHAIR:** Paul Albert, National Cancer Institute, National Institutes of Health

**3:45 Misclassified Group Tested Current Status Data**

Nicholas P. Jewell\*, London School of Hygiene & Tropical Medicine; Lucia Petitto, Harvard T.H. Chan School of Public Health

**4:10 Generalized Additive Regression for Group Testing Data**

Joshua M. Tebbs\*, University of South Carolina; Yan Liu, University of Nevada, Reno; Christopher S. McMahan, Clemson University; Chris R. Bilder, University of Nebraska, Lincoln

**4:35 Grouping Methods for Estimating the Prevalences of Rare Traits from Complex Survey Data that Preserve Confidentiality of Respondents**

Noorie Hyun\*, Medical College of Wisconsin; Joseph L. Gastwirth, The George Washington University; Barry I. Graubard, National Cancer Institute, National Institutes of Health

**5:00 Nonparametric Estimation of a Continuous Distribution Following Group Testing**

Aiyi Liu\* and Wei Zhang, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health; Qizhai Li, Chinese Academy of Sciences; Paul S. Albert, National Cancer Institute, National Institutes of Health



# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 5:25 Floor Discussion

### 89. ADAPTIVE AND BAYESIAN ADAPTIVE DESIGN IN BIOEQUIVALENCE AND BIOSIMILAR STUDIES

*Franklin Hall 13, 4th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section

**ORGANIZER:** Haiwen Shi, U.S. Food and Drug Administration

**CHAIR:** Haiwen Shi, U.S. Food and Drug Administration

### 3:45 Optimal Adaptive Sequential Designs for Crossover Bioequivalence Studies

Donald J. Schuurmann\*, U.S. Food and Drug Administration

### 4:15 A Bayesian Adaptive Design for Biosimilar Clinical Trials Using Calibrated Power Prior

Ying Yuan\*, University of Texas MD Anderson Cancer Center

### 4:45 Sequential Bioequivalence

A. Lawrence Gould\*, Merck Research Laboratories

### 5:15 Discussant:

Veronica Taylor, U.S. Food and Drug Administration

### 90. METHODS TO ROBUSTLY INCORPORATE EXTERNAL DATA INTO GENETIC TESTS

*Salon C, 5th Floor*

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Audrey Hendricks, University of Colorado Denver

**CHAIR:** Rhonda Bacher, University of Florida

### 3:45 Ancestry-Matched Allele Frequency Estimates

Tracy Ke\*, Harvard University; Alex Bloemendal, Danfeng Chen, Claire Churchhouse, Benjamin Neale and Duncan Palmer, Broad Institute; Klea Panayidou, Carnegie Mellon University; Katherine Tashman, Broad Institute; Kathryn Roeder, Carnegie Mellon University

### 4:10 Empowering External Multi-Ethnic Data in Modern, Diverse Studies

Chris Gignoux\*, Colorado Center for Personalized Medicine

### 4:35 Integrating External Controls to Association Tests

Seunggeun Lee\*, University of Michigan

### 5:00 ProxECAT: A Case-Control Gene Region Association Test using Allele Frequencies from Public Controls

Audrey E. Hendricks\*, University of Colorado Denver

## 5:25 Floor Discussion

### 91. DEVELOPING COLLABORATIVE SKILLS FOR SUCCESSFUL CAREERS IN BIOSTATISTICS AND DATA SCIENCE

*Franklin Hall 2, 4th Floor*

**SPONSOR:** CENS, ENAR

**ORGANIZERS:** Kylie Ainslie, Imperial College London and Jing Li, Indiana University

**CHAIR:** Alex Kaizer, University of Colorado

### 3:45 Panel Discussion:

Lei Shen, Eli Lilly and Company  
Patrick Staples, Mindstrong Health  
Eric Ross, Fox Chase Cancer Center  
Barret Schloerke, RStudio

## 5:15 Floor Discussion

### 92. NEW APPROACHES TO CAUSAL INFERENCE UNDER INTERFERENCE: BRINGING METHODOLOGICAL INNOVATIONS INTO PRACTICE

*Salon D, 5th Floor*

**SPONSOR:** IMS

**ORGANIZER:** Xiaoxuan Cai, Yale University

**CHAIR:** Forrest Crawford, Yale University

### 3:45 Design and Analysis of Vaccine Studies in the Presence of Interference

M. Elizabeth Halloran\*, Fred Hutchinson Cancer Research Center and University of Washington

### 4:10 Nonparametric Identification of Causal Intervention Effects under Contagion

Wen Wei Loh\*, Ghent University; Forrest W. Crawford, Yale University

### 4:35 Pairwise Regression in Infectious Disease Epidemiology with Applications to Ebola and Cholera

Eben Kenah\*, The Ohio State University

### 5:00 New Approaches to Causal Inference under Interference: Bringing Methodological Innovations into Practice

Xiaoxuan Cai\*, Yale University; M. Elizabeth Halloran, Fred Hutchinson Cancer Research Center and University of Washington; Wen Wei Loh, Ghent University; Eben Kenah, The Ohio State University; Forrest W. Crawford, Yale University

## 5:25 Floor Discussion

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 93. CONTRIBUTED PAPERS: DESIGN AND ANALYSIS OF CLINICAL TRIALS

*Franklin Hall 3, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Emily Slade, University of Kentucky

- 3:45** **A Unified Approach for Frequentist and Bayesian Hypothesis Testing in Two-Arm Fixed-Sample Clinical Trials with Binary Outcomes**  
Zhenning Yu\*, Viswanathan Ramakrishnan and Caitlyn Meinzer, Medical University of South Carolina

- 4:00** **Designing Two Arm Clinical Trials with Historical Data Using BayesCTDesign**  
Barry S. Eggleston\* and Diane J. Catellier, RTI International; Joseph G. Ibrahim, University of North Carolina, Chapel Hill

- 4:15** **Randomization Inference for a Treatment Effect in Cluster Randomized Trials**  
Dustin J. Rabideau\*, Harvard T. H. Chan School of Public Health; Rui Wang, Harvard T. H. Chan School of Public Health, Harvard Medical School and Harvard Pilgrim Health Care Institute

- 4:30** **Is Correcting for Multiple Testing in a Platform Trial Necessary?**  
Jessica R. Overbey\*, Mailman School of Public Health, Columbia University and Icahn School of Medicine at Mount Sinai; Ying Kuen K. Cheung, Mailman School of Public Health, Columbia University; Emilia Bagiella, Icahn School of Medicine at Mount Sinai

- 4:45** **Sequential Event Rate Monitoring in Clinical Trials**  
Dong-Yun Kim\*, National Heart, Lung, and Blood Institute, National Institutes of Health; Sung-Min Han, Open Source Electronic Health Record Alliance (OSEHRA)

- 5:00** **Group Sequential Enrichment Designs Based on Adaptive Regression of Response and Survival Time on High Dimensional Covariates**  
Yeonhee Park\*, Medical University of South Carolina; Suyu Liu, Peter Thall and Ying Yuan, University of Texas MD Anderson Cancer Center

- 5:15** **Floor Discussion**

## 94. CONTRIBUTED PAPERS: SEMIPARAMETRIC, NONPARAMETRIC, AND EMPIRICAL LIKELIHOOD MODELS

*405, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Inyoung Kim, Virginia Tech University

- 3:45** **The Behrens-Fisher Problem in General Factorial Designs with Covariates**  
Cong Cao\* and Frank Konietzschke, University of Texas, Dallas

- 4:00** **Systems of Partially Linear Models (SPLM) for Multi-Center Studies**  
Lei Yang\* and Yongzhao Shao, New York University

- 4:15** **Comparison of Two Transformation Models**  
Yuqi Tian\* and Bryan E. Shepherd, Vanderbilt University; Chun Li, Case Western Reserve University; Torsten Hothorn, University of Zurich; Frank E. Harrell, Vanderbilt University

- 4:30** **On Externally Calibrating Time-Dependent Absolute Risk for Time-to-Event Outcome**  
Jiayin Zheng\*, Li Hsu and Yingye Zheng, Fred Hutchinson Cancer Research Center

- 4:45** **A General Information Criterion for Model Selection based on Empirical Likelihood**  
Chixiang Chen\*, Rongling Wu and Ming Wang, The Pennsylvania State University

- 5:00** **Adjusting for Participation Bias in Case-Control Genetic Association Studies with Genotype Data Supplemented from Family Members: An Empirical Likelihood-based Estimating Equation Approach**  
Le Wang\*, Villanova University; Zhengbang Li, Central China Normal University; Clarice Weinberg, National Institute of Environmental Health Sciences, National Institutes of Health; Jinbo Chen, University of Pennsylvania

- 5:15** **Floor Discussion**

## 95. CONTRIBUTED PAPERS: BAYESIAN APPROACHES TO HIGH DIMENSIONAL DATA

*Franklin Hall 4, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Arman Oganisian, University of Pennsylvania

- 3:45** **Latent Mixtures of Functions to Characterize the Complex Exposure Relationships of Pesticides on Cancer Incidence**  
Sung Duk Kim\* and Paul S. Albert, National Cancer Institute, National Institutes of Health

- 4:00** **Prior Knowledge Guided Ultrahigh-Dimensional Variable Screening**  
Jie He\* and Jian Kang, University of Michigan

- 4:15** **Semiparametric Bayesian Kernel Survival Model for Highly Correlated High-Dimensional Data**  
Lin Zhang\*, Eli Lilly and Company (previously at Virginia Tech University); Inyoung Kim, Virginia Tech University

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

**4:30 Bayesian Variable Selection in High-Dimensional EEG Data Using Spatial Structured Spike and Slab Prior**  
Shariq Mohammed\*, University of Michigan; Dipak Kumar Dey and Yuping Zhang, University of Connecticut

**4:45 Weighted Dirichlet Process Modeling for Functional Clustering with Application in Matched Case-Crossover Studies**  
Wenyu Gao\* and Inyoung Kim, Virginia Tech University

**5:00 Floor Discussion**

## 96. CONTRIBUTED PAPERS: FUNCTIONAL DATA ANALYSIS METHODS

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Josh Barback, Harvard T. H. Chan School of Public Health

**3:45 Probabilistic K-Mean with Local Alignment for Functional Motif Discovery**  
Marzia A. Cremona\* and Francesca Chiaromonte, The Pennsylvania State University

**4:00 A Decomposable Model for Analyzing Multivariate Functional Data**  
Luo Xiao\*, North Carolina State University

**4:15 Model Testing for Generalized Scalar-on-Function Linear Models**  
Stephanie T. Chen\*, Luo Xiao and Ana-Maria Staicu, North Carolina State University

**4:30 Regression Analyses of Distributions using Quantile Functional Regression**  
Hojin Yang\*, University of Texas MD Anderson Cancer Center; Veerabhadran Baladandayuthapani, University of Michigan; Jeffrey S. Morris, University of Texas MD Anderson Cancer Center

**4:45 A Bayesian Model for Classification and Selection of Functional Predictors using Longitudinal MRI Data from ADNI**  
Asish K. Banik\* and Taps Maiti, Michigan State University

**5:00 Multilevel Localized-Variate PCA for Clustered Multivariate Functional Data**  
Jun Zhang\*, Greg J. Siegle and Robert T. Krafty, University of Pittsburgh

**5:15 Floor Discussion**

## 97. CONTRIBUTED PAPERS: NEXT GENERATION SEQUENCING

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Nicholas Lytal, University of Arizona

**3:45 A Probabilistic Model to Estimate the Temporal Order of Pathway Mutations During Tumorigenesis**  
Menghan Wang\*, Chunming Liu, Chi Wang and Arnold Stromberg, University of Kentucky

**4:00 Probabilistic Index Models for Testing Differential Gene Expression in Single-Cell RNA Sequencing (scRNA-seq) Data**  
Alemu Takele Assefa\*, Olivier Thas and Jo Vandesompele, Ghent University, Belgium

**4:15 Detection of Differentially Expressed Genes in Discrete Single-Cell RNA Sequencing Data Using a Hurdle Model with Correlated Random Effects**  
Michael Sekula\* and Jeremy Gaskins, University of Louisville; Susmita Datta, University of Florida

**4:30 Transfer Learning for Clustering Analysis from Single-Cell RNA-seq Data**  
Jian Hu\*, University of Pennsylvania, Perelman School of Medicine; Xiangjie Li, Renmin University of China; Gang Hu, Nankai University; Mingyao Liu, University of Pennsylvania, Perelman School of Medicine

**4:45 Incorporating Single-Cell RNA-seq Data to Infer Allele-Specific Expression**  
Jiaxin Fan\*, Rui Xiao and Mingyao Li, University of Pennsylvania, Perelman School of Medicine

**5:00 A Minimax Optimal Test for Rare-Variant Analysis in Whole-genome Sequencing Studies**  
Yaowu Liu\* and Xihong Lin, Harvard University

**5:15 BAMM-SC: A Bayesian Mixture Model for Clustering Droplet-Based Single Cell Transcriptomic Data from Population Studies**  
Zhe Sun\* and Ying Ding, University of Pittsburgh; Wei Chen, Children's Hospital of Pittsburgh of UPMC, University of Pittsburgh; Ming Hu, Cleveland Clinic Foundation

Denotes student award winner

# SCIENTIFIC PROGRAM

TUESDAY, MARCH 26

## 98. CONTRIBUTED PAPERS: COMPETING RISKS AND CURE MODELS

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Tao Sun, University of Pittsburgh

## ENAR BUSINESS MEETING AND SPONSOR/ EXHIBITOR MIXER

5:45 P.M.—7:00 P.M.

Franklin Hall Foyer, 4th Floor

### 3:45 General Regression Model for the Subdistribution of a Competing Risk under Left-Truncation and Right-Censoring

Anna Bellach\*, University of Washington; Michael R. Kosorok, University of North Carolina, Chapel Hill; Peter Gilbert, Fred Hutchinson Cancer Research Center; Jason P. Fine, University of North Carolina, Chapel Hill

### 4:00 Doubly Robust Outcome Weighted Learning Estimator for Competing Risk Data with Group Variable Selection

Yizeng He\*, Medical College of Wisconsin; Mi-Ok Kim, University of California, San Francisco; Soyoung Kim and Kwang Woo Ahn, Medical College of Wisconsin

### 4:15 Covariate Adjustment for Treatment Effect on Competing Risks Data in Randomized Clinical Trials

Youngjoo Cho\* and Cheng Zheng, University of Wisconsin, Milwaukee; Meijie Zhang, Medical College of Wisconsin

### 4:30 Marginal Cure Rate Models for Long-Term Survivors

Jianfeng Chen\* and Wei-Wen Hsu, Kansas State University; David Todem, Michigan State University; KyungMann Kim, University of Wisconsin, Madison

### 4:45 An Application of the Cure Model to a Cardiovascular Clinical Trial

Varadan V. Sevilimedu\*, Department of Veteran Affairs and Yale University; Shuangge Ma, Yale University; Pamela Hartigan and Tassos C. Kyriakides, Department of Veteran Affairs

### 5:00 Estimation the Serial Interval Using Cure Models

Laura F. White\*, Helen E. Jenkins, Paola Sebastiani and Yicheng Ma, Boston University

### 5:15 Floor Discussion



# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

WEDNESDAY, MARCH 27

8:30–10:15 A.M.

## 99. MONITORING HEALTH BEHAVIORS WITH MULTI-SENSOR MOBILE TECHNOLOGY

Salon B, 5th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Mental Health Statistics Section

**ORGANIZER:** Vadim Zipunnikov, Johns Hopkins Bloomberg School of Public Health

**CHAIR:** Jiawei Bai, Johns Hopkins University

### 8:30 Variable Selection in the Concurrent Functional Linear Model

Jeff Goldsmith\*, Columbia University; Joseph E. Schwartz, Columbia University Medical Center and Stony Brook University

### 8:55 Statistical Modelling of Cross-Systems Biomarkers

Vadim Zipunnikov\*, Johns Hopkins Bloomberg School of Public Health; Haochang Shou, University of Pennsylvania; Mike Xiao and Kathleen Merikangas, National Institute of Mental Health, National Institutes of Health

### 9:20 Translational Biomarkers for Quality of Sleep

Dmitri Volfson\*, Takeda Pharmaceutical Company; Brian Tracey, Tufts University; Tamas Kiss, Hungarian Academy of Sciences; Derek Buhl, Takeda Pharmaceutical Company

### 9:45 Statistical Modeling for Integrating Data from Multiple Wearable Sensors to Detect Affect Lability

Fengqing Zhang\*, Tinashe Tapera and Adrienne Juarascio, Drexel University

### 10:10 Floor Discussion

## 100. CURRENT METHODS TO ADDRESS DATA ERRORS IN ELECTRONIC HEALTH RECORDS

Salon C, 5th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistics in Epidemiology Section

**ORGANIZER:** Yong Chen, University of Pennsylvania

**CHAIR:** Pamela Shaw, University of Pennsylvania

### 8:30 Multiple Imputation to Address Data Errors in Electronic Health Record Analyses: Advantages and Disadvantages

Bryan E. Shepherd\*, Vanderbilt University; Mark J. Giganti, Harvard University

### 9:00 Raking and Regression Calibration: Methods to Address Bias Induced from Correlated Covariate and Time-to-Event Error

Eric J. Oh\* and Pamela A. Shaw, University of Pennsylvania

### 9:30 An Augmented Estimation Procedure for EHR-based Association Studies Accounting for Differential Misclassification

Yong Chen\*, Jiayi Tong and Jing Huang, University of Pennsylvania; Xuan Wang, Zhejiang University; Jessica Chubak, Kaiser Permanente Washington Health Research Institute; Rebecca Hubbard, University of Pennsylvania

### 10:00 Discussant:

Thomas Lumley, University of Auckland

## 101. FINDING THE RIGHT ACADEMIC FIT: EXPERIENCES FROM FACULTY ACROSS THE ACADEMIC SPECTRUM

Franklin Hall 1, 4th Floor

**SPONSOR:** ENAR, CENS

**ORGANIZER:** Leslie McClure, Dornsife School of Public Health at Drexel University

**CHAIR:** John Muschelli, Johns Hopkins Bloomberg School of Public Health

### 8:30 Panel Discussion:

Jianwen Cai, University of North Carolina, Chapel Hill  
Alexandra Hanlon, University of Pennsylvania  
Leslie McClure, Dornsife School of Public Health at Drexel University  
Sujata Patil, Memorial Sloan Kettering Cancer Center  
Randall H. Rieger, West Chester University

### 10:10 Floor Discussion

## 102. NOVEL INTEGRATIVE OMICS APPROACHES FOR UNDERSTANDING COMPLEX HUMAN DISEASES

Franklin Hall 2, 4th Floor

**SPONSOR:** ENAR, ASA Statistics in Genomics and Genetics Section

**ORGANIZERS:** Ran Tao and Xue Zhong, Vanderbilt University

**CHAIR:** Ran Tao, Vanderbilt University

### 8:30 Integrative Analysis of Incomplete Multi-Omics Data

Danyu Lin\*, University of North Carolina, Chapel Hill

### 8:55 An Integrative Framework to Empower Genomics-Informed Analysis of Whole Genome Sequencing Data for Complex Diseases

Bingshan Li\*, Vanderbilt University

### 9:20 Probabilistic Two Sample Mendelian Randomization for Genome-Wide Association Studies

Xiang Zhou\* and Zhongshang Yuan, University of Michigan

### 9:45 A Semi-Supervised Approach for Predicting Cell-Type Specific Functional Consequences of Non-Coding Variation using MPRA

Zihuai He\*, Stanford University; Linxi Liu, Columbia University; Kai Wang, Raymond G. Perelman Center for Cellular and Molecular Therapeutics, Children's Hospital of Philadelphia; Iuliana Ionita-Laza, Columbia University

### 10:10 Floor Discussion

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

## 103. TEACHING DATA SCIENCE THROUGH CASE-STUDIES

Salon D, 5th Floor

**SPONSOR:** ASA Statistical Learning and Data Science Section

**ORGANIZER:** Stephanie Hicks, Johns Hopkins Bloomberg School of Public Health

**CHAIR:** Stephanie Hicks, Johns Hopkins Bloomberg School of Public Health

### 8:30 Motivating Data Science Through Case Studies in Public Health

Leah R. Jager\*, Johns Hopkins Bloomberg School of Public Health

### 8:55 Teaching Genomic Data Science: Summarization, Exploration, and Reproducibility

Michael I. Love\*, University of North Carolina, Chapel Hill

### 9:20 Before Teaching Data Science, Let's First Understand How People Do It

Rebecca Nugent\*, Philipp Burckhardt and Ronald Yurko, Carnegie Mellon Statistics & Data Science

### 9:45 Introduction to Data Science, Case-by-Case

Mine Cetinkaya-Rundel\*, Duke University and RStudio

### 10:10 Floor Discussion

## 104. NONCONVEX OPTIMIZATION AND BIOLOGICAL APPLICATIONS

Franklin Hall 13, 4th Floor

**SPONSOR:** IMS

**ORGANIZER:** Benjamin Risk, Emory University

**CHAIR:** Irina Gaynanova, Texas A&M University

### 8:30 Local False Discovery Rates for Nonconvex Penalties in High-Dimensional Regression Models

Patrick Breheny\* and Ryan E. Miller, University of Iowa

### 8:55 It's just a Matter of Perspective - Robust Regression for Microbiome Data via Perspective M-estimation

Christian L. Mueller\*, Flatiron Institute, Simons Foundation

### 9:20 Relax and Split Algorithm for ICA

Peng Zheng\*, University of Washington; Benjamin Risk, Emory University; Irina Gaynanova, Texas A&M University; Aleksandr Y. Aravkin, University of Washington

### 9:45 Integrated Principal Component Analysis

Genevera I. Allen\*, Rice University and Baylor College of Medicine; Tiffany M. Tang, University of California, Berkeley

### 10:10 Floor Discussion

## 105. CONTRIBUTED PAPERS: BIOPHARMACEUTICAL RESEARCH AND CLINICAL TRIALS

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Jiayin Zheng, Fred Hutchinson Cancer Research Center

### 8:30 In Silico Clinical Trial Simulation and Virtual Patients' Generation

Philippe Saint Pierre\* and Nicolas J. Savy, Toulouse Institute of Mathematics

### 8:45 Nonparametric Tests for Transition Probabilities in Markov Multi-State Models

Giorgos Bakoyannis\*, Indiana University

### 9:00 Trigger Strategy in Repeated Tests on Multiple Hypotheses

Jiangtao Gou\*, Fox Chase Cancer Center, Temple University Health System

### 9:15 MCP-Mod for Exposure-Response Information

Gustavo Amorim\*, Vanderbilt University Medical Center; An Vandebosch, Jose Pinheiro, Joris Menten and Kim Stuyckens, Janssen Pharmaceutica

### 9:30 Evaluating the Finite Sample Properties of Baseline Covariate Adjustment in Randomized Trials: Application to Time to Event and Binary Outcomes

Su Jin Lim\*, Johns Hopkins University School of Medicine; Elizabeth Colantuoni, Johns Hopkins Bloomberg School of Public Health

### 9:45 Floor Discussion

## 106. CONTRIBUTED PAPERS: MISSING DATA

405, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Dong-Yun Kim, National Heart Lung and Blood Institute, National Institutes of Health

### 8:30 A Doubly-Robust Method to Handle Missing Multilevel Outcome Data with Application to the China Health and Nutrition Survey

Nicole M. Butera\*, Donglin Zeng, Annie Green Howard, Penny Gordon-Larsen and Jianwen Cai, University of North Carolina, Chapel Hill

### 8:45 Reproducibility of High Throughput Experiments in Case of Missing Data

Roopali Singh\*, Feipeng Zhang and Qunhua Li, The Pennsylvania State University

### 9:00 Multiple Imputation Strategies for Handling Missing Data When Generalizing Randomized Clinical Trial Findings Through Propensity Score-Based Methodologies

Albee Ling\*, Maya Mathur, Kris Kapphahn, Maria Montez-Rath and Manisha Desai, Stanford University

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

**9:15 Proper Specification of MICE Imputation Models for Data with Interactions: New Developments and Practical Recommendations for R Users**

Emily Slade\*, University of Kentucky

**9:30 Variance Estimation When Combining Inverse Probability Weighting and Multiple Imputation in Electronic Health Records-Based Research**

Tanayott Thaweethai\*, Harvard University; Sebastien Haneuse, Harvard T.H. Chan School of Public Health; David Arterburn, Kaiser Permanente Washington Health Research Institute

**9:45 Model-Based Phenotyping in Electronic Health Records with Data for Anchor-labeled Cases and Unlabeled Patients**

Lingjiao Zhang\* and Xiruo Ding, University of Pennsylvania; Yanyuan Ma, The Pennsylvania State University; Naveen Muthu, Jason Moore, Daniel Herman and Jinbo Chen, University of Pennsylvania, Philadelphia

**10:00 Fully Bayesian Imputation Model for Non-Random Missing Data in qPCR**

Valeriia Sherina\*, Matthew N. McCall and Tanzy M. T. Love, University of Rochester Medical Center

## 107. CONTRIBUTED PAPERS: BAYESIAN COMPUTATIONAL AND MODELING METHODS

*Franklin Hall 4, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Wenli Sun, University of Pennsylvania

**8:30 Sampling Prudently using Inversion Spheres (SPInS) on the Simplex**

Sharang Chaudhry\*, Daniel Lautzenheiser and Kaushik Ghosh, University of Nevada Las Vegas

**8:45 BayesMetab: Bayesian Modelling Approach in Treating Missing Values in Metabolomic Studies**

Jasmit Shah\*, Aga Khan University; Guy N. Brock, The Ohio State University; Jeremy Gaskins, University of Louisville

**9:00 A Bayesian Markov Model for Personalized Benefit-Risk Assessment**

Dongyan Yan\*, University of Missouri; Subharup Guha, University of Florida; Chul Ahn and Ram Tiwari, U.S. Food and Drug Administration

**9:15 Iterated Multi-Source Exchangeability Models for Individualized Inference with an Application to Mobile Sensor Data**

Roland Z. Brown\* and Julian Wolfson, University of Minnesota

**9:30 A Novel Bayesian Predictive Modelling in Time-to-Event Analysis using Multiple-Imputation Techniques**

Zhe (Vincent) Chen\* and Kalyanee Viraswami-Appanna, Novartis Pharmaceuticals Corporation

**9:45 A Latent Class Based Joint Model for Recurrence and Termination: A Bayesian Recourse**

Zhixing Xu\*, Debjayoti Sinha and Jonathan Bradley, Florida State University

**10:00 Floor Discussion**

## 108. CONTRIBUTED PAPERS: CAUSAL EFFECT MODELING (MEDIATION/VARIABLE SELECTION/LONGITUDINAL)

*406, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Caroline Groth, Northwestern University Feinberg School of Medicine

**8:30 The Role of Body Mass Index at Diagnosis on Black-White Disparities in Colorectal Cancer Survival: A Density Regression Mediation Approach**

Katrina L. Devick\*, Harvard T.H. Chan School of Public Health; Linda Valeri, Columbia Mailman School of Public Health; Jarvis Chen, Harvard T.H. Chan School of Public Health; Alejandro Jara, Pontificia Universidad Católica de Chile; Marie-Abèle Bind, Harvard University; Brent A. Coull, Harvard T.H. Chan School of Public Health

**8:45 Unified Mediation Analysis Approach to Complex Data of Mixed Types via Copula Models**

Wei Hao\* and Peter X.K. Song, University of Michigan

**9:00 Estimating Causal Mediation Effects from a Single Regression Model**

Christina T. Saunders\* and Jeffrey D. Blume, Vanderbilt University

**9:15 Mediator Selection via the Lasso with Nonparametric Confounding Control**

Jeremiah Jones\* and Ashkan Ertefaie, University of Rochester

**9:30 Estimating Time-Varying Causal Effect Moderation in Mobile Health with Binary Outcomes**

Tianchen Qian\*, Harvard University; Hyesun Yoo, Predrag Klasnja and Daniel Almirall, University of Michigan; Susan A. Murphy, Harvard University

**9:45 Estimating Causal Effects with Longitudinal Data in a Bayesian Framework**

Kuan Liu\* and Olli Saarela, University of Toronto; Eleanor Pullenayegum, University of Toronto, The Hospital for Sick Children

Denotes student award winner

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

## 10:00 Brand vs. Generic: Addressing Non-Adherence, Secular Trends, and Non-Overlap

Lamar Hunt\* and Irene B. Murimi, Johns Hopkins Bloomberg School of Public Health and OptumLabs Visiting Fellows; Daniel O. Scharfstein, Johns Hopkins Bloomberg School of Public Health; Jodi B. Segal, Johns Hopkins School of Medicine; Marissa J. Seamans, Johns Hopkins Bloomberg School of Public Health; Ravi Varadhan, Johns Hopkins Center on Aging and Health

## 109. CONTRIBUTED PAPERS: MICROBIOME DATA ANALYSIS WITH ZERO INFLATION AND/OR MODEL SELECTION

Franklin Hall 3, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Hyunwook Koh, Johns Hopkins Bloomberg School of Public Health

## 8:30 An Integrative Bayesian Zero-Inflated Negative Binomial Model for Microbiome Data Analysis

Shuang Jiang\*, Southern Methodist University; Guanghua Xiao, Andrew Y. Koh, Yang Xie, Qiwei Li and Xiaowei Zhan, University of Texas Southwestern Medical Center

## 8:45 Bayesian Hierarchical Zero-Inflated Negative Binomial Models with Applications to High-Dimensional Human Microbiome Count Data

Amanda H. Pendegraft\* and Nengjun Yi, University of Alabama at Birmingham

## 9:00 Model Selection for Longitudinal Microbiome Data with Excess Zeros

Tony A. Chen\*, Princeton University; Yilun Sun, Hana Hakim, Ronald Dallas, Jason Rosch, Sima Jeha and Li Tang, St. Jude Children's Research Hospital

## 9:15 Bayesian Variable Selection in Regression with Compositional Covariates

Liangliang Zhang\*, University of Texas MD Anderson Cancer Center

## 9:30 Compositional Knockoff Filter for FDR Control in Microbiome Regression Analysis

Arun A. Srinivasan\*, Lingzhou Xue and Xiang Zhan, The Pennsylvania State University

## 9:45 Generalized Biplots for the Analysis of Human Microbiome

Yue Wang\* and Timothy W. Randolph, Fred Hutchinson Cancer Research Center; Ali Shojaie, University of Washington; Jing Ma, Fred Hutchinson Cancer Research Center

## 10:00 Floor Discussion

## 110. CONTRIBUTED PAPERS: RECURRENT EVENTS OR MULTIPLE TIME-TO-EVENT DATA

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Bo Hu, Columbia University

## 8:30 Regression Analysis of Recurrent Event Data with Measurement Error

Yixin Ren\* and Xin He, University of Maryland, College Park

## 8:45 A General Class of Semiparametric Models for Biased Recurrent Event Data

Russell S. Stocker\*, Indiana University of Pennsylvania; Akim Adekpedjou, Missouri University of Science and Technology

## 9:00 Penalized Survival Models for the Analysis of Alternating Recurrent Event Data

Lili Wang\*, Kevin He and Douglas E. Schaubel, University of Michigan

## 9:15 A Time-Varying Joint Frailty-Copula Model for Analyzing Recurrent Events and a Terminal Event: An Application to the Cardiovascular Health Study

Zheng Li\*, Novartis Pharmaceuticals Corporation; Vernon M. Chinchilli and Ming Wang, The Pennsylvania State University

## 9:30 Dynamic Regression with Recurrent Events

Jae Eui Soh\* and Yijian Huang, Emory University

## 9:45 Spearman's Correlation for Estimating the Association between Two Time-to-Event Outcomes

Svetlana K. Eden\*, Vanderbilt University; Chun Li, Case Western Reserve University; Bryan E. Shepherd, Vanderbilt University

## 10:00 Floor Discussion

## WEDNESDAY, MARCH 27

10:15–10:30 A.M.

## REFRESHMENT BREAK WITH OUR EXHIBITORS

Franklin Hall Foyer, 4th Floor

## WEDNESDAY, MARCH 27

10:30 A.M.—12:15 P.M.

## 111. INDIVIDUALIZED EVIDENCE FOR MEDICAL DECISION MAKING: PRINCIPLES AND PRACTICES

Franklin Hall 1, 4th Floor

**SPONSOR:** ASA Section: Medical Devices and Diagnostics

**ORGANIZER:** Gene Pennello, U.S. Food and Drug Administration

**CHAIR:** Qin Li, U.S. Food and Drug Administration

Denotes student award winner



# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

## 10:30 Bayesian Hierarchical Models for Individualized Health

Scott L. Zeger\*, Johns Hopkins University

## 10:55 Personalized Bayesian Minimum-Risk Decisions for Treatment of Coronary Artery Disease

Laura A. Hatfield\*, Harvard Medical School

## 11:20 Assessing Potential Clinical Impact with Net Benefit Measures

Tracey L. Marsh\*, Fred Hutchinson Cancer Research Center

## 11:45 Individualized Evidence for Medical Decision Making: Principles and Practices

David M. Kent\*, Tufts Medical Center

## 12:10 Floor Discussion

### 112. SOME NEW PERSPECTIVES AND DEVELOPMENTS FOR DATA INTEGRATION IN THE ERA OF DATA SCIENCE

Franklin Hall 2, 4th Floor

**SPONSOR:** ENAR, ASA Biometrics Section, ASA Statistical Learning and Data Science Section

**ORGANIZER:** Peisong Han, University of Michigan

**CHAIR:** Peisong Han, University of Michigan

## 10:30 Using Synthetic Data to Update an Established Prediction Model with New Biomarkers

Jeremy Taylor\*, Tian Gu and Bhramar Mukherjee, University of Michigan

## 10:55 Integrative Data Analytics and Confederate Inference

Peter XK Song\*, University of Michigan; Lu Tang, University of Pittsburgh; Ling Zhou, University of Michigan

## 11:20 Generalized Meta-Analysis for Data Integration with Summary-Level Data

Nilanjan Chatterjee\*, Prosenjit Kundu and Runlong Tang, Johns Hopkins University

## 11:45 A Superpopulation Approach to Case-Control Studies

Yanyuan Ma\*, The Pennsylvania State University

## 12:10 Floor Discussion

### 113. BAYESIAN METHODS FOR SPATIAL AND SPATIO-TEMPORAL MODELING OF HEALTH DATA

Salon B, 5th Floor

**SPONSOR:** ENAR

**ORGANIZER:** Jing Wu, University of Rhode Island

**CHAIR:** Jing Wu, University of Rhode Island

## 10:30 Step Change Detection and Forecasting of Vector-Borne Diseases

Gavino Puggioni\* and Jing Wu, University of Rhode Island

## 10:55 Bayesian Disaggregation of Spatio-Temporal Community Indicators Estimated via Surveys: An Application to the American Community Survey

Veronica J. Berrocal\* and Marco H. Benedetti, University of Michigan

## 11:20 Age-Specific Distributed Lag Models for Heat-Related Mortality

Matthew J. Heaton\*, Brigham Young University; Cassandra Olenick and Olga V. Wilhelmi, National Center for Atmospheric Research

## 11:45 Restricted Nonparametric Mixtures Models for Disease Clustering

Abel Rodriguez\* and Claudia Wehrhahn, University of California, Santa Cruz

## 12:10 Floor Discussion

### 114. RECENT ADVANCES IN CAUSAL INFERENCE FOR SURVIVAL ANALYSIS

Salon C, 5th Floor

**SPONSOR:** ASA Biometrics Section, ASA Statistics in Epidemiology Section, ASA Health Policy Statistics Section, ASA Statistical Learning and Data Science Section, ASA Mental Health Statistics Section

**ORGANIZERS:** Eric Tchetgen Tchetgen, University of Pennsylvania and Linbo Wang, University of Toronto

**CHAIR:** Eric Tchetgen Tchetgen, University of Pennsylvania

## 10:30 Adjusting for Time-Varying Confounders in Survival Analysis using Structural Nested Cumulative Survival Time Models

Stijn Vansteelandt\*, Ghent University and London School of Hygiene and Tropical Medicine; Shaun Seaman, Cambridge University; Oliver Dukes, Ghent University; Ruth Keogh, London School of Hygiene and Tropical Medicine

## 11:00 Instrumental Variables Estimation with Competing Risk Data

Torben Martinussen\*, University of Copenhagen; Stijn Vansteelandt, Ghent University

## 11:30 Instrumental Variable Estimation of a Cox Marginal Structural Model with Endogenous Time-Varying Exposure

Yifan Cui\*, Haben Michael, and Eric Tchetgen Tchetgen, The Wharton School, University of Pennsylvania

## 12:00 Discussant:

Ilya Shpitser, Johns Hopkins University

### 115. NOVEL STATISTICAL METHODS FOR ANALYSIS OF MICROBIOME DATA

Salon D, 5th Floor

**SPONSOR:** ASA Statistics in Genomics and Genetics Section

**ORGANIZER:** Xiang Zhan, The Pennsylvania State University

**CHAIR:** Anna Plantinga, Williams College

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

**10:30 High Dimensional Mediation Model for Microbial Abundance Data**  
Ni Zhao\*, Johns Hopkins University; Junxian Chen, The Hong Kong Polytechnic University

**10:55 A Sparse Causal Mediation Model for Microbiome Data Analysis**  
Huilin Li\*, Chan Wang, Jiyuan Hu and Martin Blaser, New York University

**11:20 A Robust and Powerful Framework for Microbiome Biomarker Discovery**  
Jun Chen\*, Mayo Clinic; Li Chen, Auburn University

**11:45 Omnidirectional Visualization of Competition and Cooperation in the Gut Microbiota**  
Rongling Wu\*, The Pennsylvania State University

**12:10 Floor Discussion**

## 116. NEW DEVELOPMENTS IN NONPARAMETRIC METHODS FOR COVARIATE SELECTION

*Franklin Hall 13, 4th Floor*

**SPONSOR:** IMS

**ORGANIZER:** Sherri Rose, Harvard Medical School

**CHAIR:** Ani Eloyan, Brown University

**10:30 Variable Prioritization in Black Box Statistical Methods**  
Lorin Crawford\*, Brown University

**10:55 Covariate Selection and Algorithmic Fairness for Continuous Outcomes in Health Plan Risk Adjustment**  
Sherri Rose\*, Harvard Medical School; Anna Zink, Harvard University

**11:20 Non-Parametric and Data-Driven Methods for Identifying Subpopulations Susceptible to the Health Effects of Air Pollution**  
Cole Brokamp\*, Cincinnati Children's Hospital Medical Center

**11:45 Dynamic Landmark Prediction for Mixture Data**  
Tanya P. Garcia\*, Texas A&M University; Layla Parast, RAND Corporation

**12:10 Floor Discussion**

## 117. CONTRIBUTED PAPERS: DYNAMIC TREATMENT REGIMENS AND EXPERIMENTAL DESIGN

*Franklin Hall 3, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Yimei Li, University of Pennsylvania

**10:30 Should I Stay or Should I Go: Selecting Individualized Stage Duration in a Sequential Multiple Assignment Randomized Trial (SMART)**  
Hayley M. Belli\* and Andrea B. Troxel, New York University Langone School of Medicine

**10:45 New Statistical Learning for Evaluating Nested Dynamic Treatment Regimes**  
Ming Tang\*, Lu Wang and Jeremy M.G. Taylor, University of Michigan

**11:00 Design and Analysis Issues for Estimating Transmission Probabilities in a Challenge Study**  
Sally Hunsberger\*, Michael A. Proschan, Alison Han and Matthew J. Memoli, National Institute of Allergy and Infectious Diseases, National Institutes of Health

**11:15 Discovery of Gene Regulatory Networks Using Adaptively-Selected Gene Perturbation Experiments**  
Michele S. Zemplenyi\* and Jeffrey W. Miller, Harvard University

**11:30 A Sample Size Calculation for Bayesian Analysis of Small n Sequential Multiple Assignment Randomized Trials (snSMARTs)**  
Boxian Wei\* and Thomas M. Braun, University of Michigan; Roy N. Tamura, University of South Florida; Kelley M. Kidwell, University of Michigan

**11:45 Design of Experiments for a Confirmatory Trial of Precision Medicine**  
Kim May Lee\* and James Wason, University of Cambridge

**12:00 Floor Discussion**

## 118. CONTRIBUTED PAPERS: HYPOTHESIS TESTING AND SAMPLE SIZE CALCULATION

*405, 4th Floor*

**SPONSOR:** ENAR

**CHAIR:** Xin Zhou, Harvard T. H. Chan School of Public Health

**10:30 Bayesian Nonparametric Test for Independence Between Random Vectors**  
Zichen Ma\* and Timothy E. Hanson, University of South Carolina

**10:45 Kernel Based-Hybrid Test for High-Dimensional Data**  
Inyoung Kim\*, Virginia Tech University

**11:00 A Non-Nested Hypothesis Testing Problem for Threshold Regression Models**  
Zonglin He\*, Fred Hutchinson Cancer Research Center

**11:15 Robust Bootstrap Testing for Nonlinear Effect in Small Sample with Kernel Ensemble**  
Wenying Deng\*, Jeremiah Zhe Liu and Brent Coull, Harvard T.H. Chan School of Public Health

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

**11:30 A Robust Hypothesis Test for Continuous Nonlinear Interactions in Nutrition-Environment Studies: A Cross-Validated Ensemble Approach**

Jeremiah Zhe Liu\*, Harvard T.H. Chan School of Public Health; Jane Lee, Boston Children's Hospital; Pi-i Debby Lin, Harvard University; Linda Valeri, Columbia Mailman School of Public Health; David Christiani and David Bellinger, Harvard T.H. Chan School of Public Health; Robert Wright, Icahn School of Medicine at Mount Sinai; Maitreyi Mazumdar, Boston Children's Hospital; Brent Coull, Harvard T.H. Chan School of Public Health

**11:45 A Goodness of Fit Test to Compare Lumped and Unlumped Markov Chains**

Anastasia M. Hartzes\* and Charity J. Morgan, University of Alabama at Birmingham

**12:00 Sample Size for Trials Comparing Group and Individual Treatments with Repeated Measures**

Robert J. Gallop\*, West Chester University

## 119. CONTRIBUTED PAPERS: MEASUREMENT ERROR

302-303, 3rd Floor

**SPONSOR:** ENAR

**CHAIR:** Carmen Tekwe, Texas A&M University

**10:30 Categorizing a Continuous Predictor Subject to Measurement Error**

Tianying Wang\*, Columbia University; Raymond Carroll, Texas A&M University; Betsabe Blas Achic, Universidade Federal de Pernambuco; Ya Su, University of Kentucky; Victor Kipnis and Kevin Dodd, National Cancer Institute, National Institutes of Health

**10:45 Efficient Inference for Two-Phase Designs with Response and Covariate Measurement Error**

Sarah C. Lotzpeich\*, Vanderbilt University; Bryan E. Shepherd, Vanderbilt University Medical Center; Pamela Shaw, University of Pennsylvania; Ran Tao, Vanderbilt University Medical Center

**11:00 Improving the Reproducibility of EHR-Based Association Studies for Pleiotropic Effects by Accounting for Phenotyping Errors**

Jiayi Tong\* and Ruowang Li, University of Pennsylvania; Doudou Zhou, University of Science and Technology of China; Rui Duan, Jason Moore and Yong Chen, University of Pennsylvania

**11:15 Bayesian Latent Class Regression for Measurement Error Correction in Self-Reported Dietary Intake**

Caroline P. Groth\* and David Aaby, Northwestern University Feinberg School of Medicine; Michael J. Daniels, University of Florida; Linda Van Horn and Juned Siddique, Northwestern University Feinberg School of Medicine

**11:30 Bayesian Approach for Handling Covariate Measurement Error when Estimating Population Treatment Effect**

Hwanhee Hong\*, Duke University; Juned Siddique, Northwestern University Feinberg School of Medicine; Elizabeth A. Stuart, Johns Hopkins Bloomberg School of Public Health

**11:45 Flexible Omnibus Test in 1:M Matched Case-Crossover Study with Measurement Error in Covariate**

Byung-Jun Kim\* and Inyoung Kim, Virginia Tech University

**12:00 Floor Discussion**

## 120. CONTRIBUTED PAPERS: ENVIRONMENTAL AND ECOLOGICAL APPLICATIONS

406, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Chi Hyun Lee, University of Massachusetts

**10:30 Integral Projection Models for Population in Columbian Ground Squirrel**

Kyoung Ju Kim\*, Auburn University

**10:45 A Bayesian Critical Window Variable Selection Method for Estimating the Impact of Air Pollution Exposure during Pregnancy**

Joshua L. Warren\* and Wenjing Kong, Yale University; Thomas J. Luben, United States Environmental Protection Agency; Howard H. Chang, Emory University

**11:00 Combining Air Pollution Estimates from Multiple Statistical Models Using Spatial Bayesian Ensemble Averaging**

Nancy L. Murray\* and Howard H. Chang, Emory University

**11:15 A Hierarchical Model for Estimating Exposure-Response Curves from Multiple Studies**

Joshua P. Keller\*, Colorado State University; Scott L. Zeger, Johns Hopkins University

**11:30 Robust Nonparametric Derivative Estimator**

Hamdy F. F. Mahmoud\*, Byung-Jun Kim and Inyoung Kim, Virginia Tech University

**11:45 Floor Discussion**

## 121. CONTRIBUTED PAPERS: STATISTICAL METHODS FOR HIGH DIMENSIONAL DATA

Franklin Hall 4, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Olivier Thas, Hasselt University

**10:30 Generalized Linked Matrix Factorization**

Michael J. O'Connell\*, Miami University

Denotes student award winner

# SCIENTIFIC PROGRAM

WEDNESDAY, MARCH 27

**10:45 Estimation and Inference for High Dimensional Generalized Linear Models: A Split and Smoothing Approach**  
Zhe Fei\* and Yi Li, University of Michigan

**11:00 Adaptive Sparse Estimation with Side Information**  
Trambak Banerje\*, Gourab Mukherjee and Wenguang Sun, University of Southern California

**11:15 Covariate Assisted Principal Regression for Covariance Matrix Outcomes**  
Yi Zhao\* and Bingkai Wang, Johns Hopkins Bloomberg School of Public Health; Stewart H. Mostofsky, Johns Hopkins University; Brian S. Caffo, Johns Hopkins Bloomberg School of Public Health; Xi Luo, Brown University

**11:30 Estimating T-Central Subspace via Marginal Third Moments**  
Weihang Ren\* and Xiangrong Yin, University of Kentucky

**11:45 Floor Discussion**

## 122. CONTRIBUTED PAPERS: LONGITUDINAL DATA AND JOINT MODELS OF LONGITUDINAL AND SURVIVAL DATA

407, 4th Floor

**SPONSOR:** ENAR

**CHAIR:** Russell Stocker, Indiana University of Pennsylvania

**10:30 Bayesian Joint Modeling of Nested Repeated Measure with the Presence of Informative Dropout**  
Enas Mustfa Ghulam\*, University of Cincinnati and Cincinnati Children's Hospital Medical Center; Rhonda D. Szczesniak, Cincinnati Children's Hospital Medical Center

**10:45 The Joint Modelling of Longitudinal Process and Censored Quantile Regression**  
Bo Hu\*, Ying Wei and Mary Beth Terry, Columbia University

**11:00 Joint Latent Class Trees: A Tree-Based Approach to Joint Modeling of Time-to-Event and Longitudinal Data**  
Jeffrey S. Simonoff\* and Ningshan Zhang, New York University

**11:15 Fusion Learning in Stratified Models by Penalized Generalized Estimating Equations**  
Lu Tang\*, University of Pittsburgh; Peter X.K. Song, University of Michigan

**11:30 Mixture of Linear Mixed Effects Models with Real Data Application**  
Yian Zhang\*, Lei Yang, Zhaoyin Zhu and Yongzhao Shao, New York University

**11:45 An Approximate Approach for Fitting Two-Part Mixed Effects Models to Longitudinal Semi-Continuous Data**  
Hyoyoung Choo-Wosoba\* and Paul S. Albert, National Cancer Institute, National Institutes of Health

**12:00 Monitoring Progression Towards Renal Failure: Lessons from a Large VA Cohort**  
Fridtjof Thomas\*, Oguz Akbilgic, Praveen K. Potukuchi and Keiichi Sumida, University of Tennessee Health Science Center; Csaba P. Kovesdy, Memphis VA Medical Center

Denotes student award winner



# SHORT COURSES

Short Courses Sunday, March 24, 2019

## SC1. Bayesian Inference and Clinical Trial Designs Using Historical Data

Full Day | 8:00 am – 5:00 pm

Franklin Hall 1, 4th Floor

Ming-Hui Chen, University of Connecticut

Fang Chen, SAS Institute Inc.

**Description:** Clinical trials are understandably expensive. However, similar trial data are often available from previous studies or experiments. Borrowing information from historical data can potentially help reducing trial cost and providing more accurate estimation while maintaining desirable qualities, such as control type I error and power. This short course provides a comprehensive review of Bayesian methods for borrowing historical information and proper use of these methods in Bayesian clinical trial designs. Several case studies are illustrated using software code that is explained in detail.

The course focuses on using historical data in design areas that include design of non-inferiority clinical trials, design of superiority clinical trials, methods for go/no-go decisions, sequential meta-analysis design, and joint analysis that combines the results from multiple trials. Special topics that are discussed include Monte Carlo simulation, Bayesian sample size determination, analysis of recurrent events, and frailty regression. The examples are shown using SAS, including the SAS macro language and the MCMC procedure, with a strong focus on technical details.

The course also includes an introduction of the Bayesian approach to inference (presented from a biopharmaceutical perspective) and outlines approaches in using and borrowing historical information, including variations of the power prior and meta-analytic-predictive prior.

## SC2. Big Data, Data Science and Deep Learning for Statistician

Full Day | 8:00 am – 5:00 pm

Franklin Hall 2, 4th Floor

Hui Lin, DowDuPont

Ming Li, Amazon

**Description:** The increasing volume and sophistication of data pose new challenges and needs for data science. There is a pressing need for data scientists who can bring actionable insight from the vast amount of data collected. In the past several years, deep learning has gained traction in many areas, and it becomes an essential tool in data scientist's toolbox. In this course, students will develop a clear understanding of the big data cloud platform, learn basic data manipulation, preprocess and machine learning skills, and understand the motivation and use cases of deep learning through hands-on exercises. We will also cover the "art" part of data science including the general data science project flow, common pitfalls, and soft skills to effectively communicate with business stakeholders. The course is for audiences with a statistical background. This course will prepare statisticians to be successful data scientists and deep learning scientist in various industries and business sectors.

We will use the Databricks community edition cloud platform throughout the training course to cover hands-on sessions including (1) big data platform using Spark through R sparklyr package; (2) introduction to Deep Neural Network, Convolutional Neural Network and Recurrent Neural Networks and their applications; (3) deep learning examples using TensorFlow through R keras package. The primary audiences for this course are (1) statistician in traditional industry sectors such as manufacturing, pharmaceutical, and banking; (2) statistician in government agencies; (3) statistical researchers in universities; (4) graduate students in statistics departments. The prerequisite knowledge is MS level education in statistics and entry level of R knowledge. No software installation is needed in participants' laptop and the cloud platform is easily accessed through web browsers such as Chrome or Firefox with the internet connection.

## SC3. Analysis of Medical Cost Data: Statistical and Econometric Tools

Full Day | 8:00 am – 5:00 pm

Franklin Hall 3, 4th Floor

Lei Liu, Washington University in St. Louis

Tina Shih, University of Texas MD Anderson Cancer Center

**Description:** Rapid growth in medical costs in the U.S. has been a major policy concern and one of the recurrent themes in presidential debates for decades. Medical cost data are routinely collected in billing records of hospitals and claims of health insurance plans (e.g., Medicare, Medicaid, or commercial insurance) or in national surveys (e.g., Medical Expenditure Panel Study). The wide availability of such data has motivated the development and application of state-of-the-art statistical and econometric methods. The policy relevance of medical cost estimates makes medical cost research extremely important because inaccurate statistical inferences could lead to misguided policy decisions.

This short course will summarize the up-to-date analytical methods for medical cost research. The short course will be co-taught by a biostatistician (Liu) and a health economist (Shih), who have collaborated on this topic for more than a decade. This interdisciplinary collaboration has resulted in multiple grants and numerous papers. Medical cost research has also sparked the interest of other quantitative scientists, leading to the development of a growing number of new analytical methods. Therefore, we think the timing is ripe to deliver a short course to summarize the latest methodological development from our group and other researchers to advance the knowledge in medical cost research.

Our short course will offer a comprehensive compilation of new approaches, modeling, and applications on medical cost analyses. It aims not only to synthesize the disparate literature of this fast growing field, but, in doing so, to foster new methodological development, new perspectives, new questions, and a broader understanding of medical cost research. While we intend to discuss recent methodological development in the analysis of medical cost data, materials will be presented in a way that is understandable to clinical researchers and policy analysts with moderate training in statistics and/or econometrics.

This application oriented short course is of interest to researchers who would apply up-to-date statistical tools to medical cost data. 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 medical cost analysis.

## SC4. StatTag for Connecting R, SAS, and Stata to Word: A Practical Approach to Reproducibility

Half Day | 8:00 am – 12:00 pm

Franklin Hall 4, 4th Floor

Abigail Baldrige, Northwestern University

Luke Rasmussen, Northwestern University

**Description:** Reproducibility, wherein data analysis and documentation is sufficient so that results can be recomputed or verified, is an increasingly important component of statistical practice. "Weaving" tools such as R Markdown facilitate reproducibility by combining narrative text and analysis code in one plain-text document, but are of limited use when manuscripts or reports must be generated in MS Word (e.g. due to journal requirements or client preference). To address this challenge, we have created StatTag, a free, open-source program that embeds statistical results from R, SAS, or Stata directly in Microsoft Word. StatTag is available as a Word plugin (Windows) or standalone application (Mac) that links statistical code files to Word documents. From Word, users attach one or more code files to an active document, and use the StatTag interface to "tag" selected statistical output – estimates, tables, or figures. The user instructs StatTag to insert the selected statistical output into the Word document, whereupon StatTag invokes the appropriate statistical software and places the result within the document text. Inserted results can then be updated automatically or on demand, and will retain their linkage to the code even when the document changes

hands, is redlined, or the text is copied and pasted elsewhere. The StatTag interface also allows direct user interaction with the code file; users may view, edit and re-run statistical code directly from Word. StatTag improves over other similar software in that it functions directly from Word, and it allows the usage of more than one statistical software and code.

In this short course, we will:

- introduce approaches for reproducible research with focus on data analysis and publication
- introduce StatTag, a reproducible research tool for Word with SAS, Stata and/or R
- lead a hands-on session during which participants will generate an abstract with StatTag in the software version of their choice and update their abstract through a brief peer review
- connect users to the StatTag knowledge base and summarize the information learned

This course is intended for a broad audience; prerequisites are experience preparing documents in Word and conducting analysis in any one of R, SAS, or Stata. In addition to the in-person course, participants will have access to an online course and materials before and after the conference.

### SC5. Personalized Medicine: Subgroup Identification in Clinical Trials

Half Day | 8:00 am – 12:00 pm

Franklin Hall 13, 4th Floor

Ilya Lipkovich, Eli Lilly

Alex Dmitrienko, Mediana, Inc

**Description:** This short course will provide a description of a broad class of statistical methods dealing with exploratory subgroup analysis in clinical trials as one of the key components of personalized medicine. This includes subgroup search/biomarker discovery methods that can be applied both in early and late-phase clinical trials. Subgroup identification from observational data will not be considered.

We will begin with a broad review of existing approaches to subgroup/biomarker identification in the context of personalized medicine illustrating the key elements of principled data-driven subgroup evaluation and then focus on a recursive partitioning method SIDES (Subgroup Identification Based on Differential Effect Search, Lipkovich et al., 2011) and its extensions SIDEScreen (Lipkovich and Dmitrienko, 2014) and Stochastic SIDEScreen (Lipkovich et al., 2017).

Key elements of SIDES and related methods will be discussed including generation of multiple promising subgroups based on different splitting criteria, evaluation of variable importance (VI), implementing VI-based biomarker screening, and addressing Type I error rate and subgroup effect inflation using resampling based methods.

Case studies from both early and late clinical development programs will be used to illustrate the principles and statistical methods introduced in this course. A software tool implementing SIDES and related methods will be presented (RSIDES package developed by the authors, <http://biopharmnet.com/subgroup-analysis-software/>).

### SC6. Design of Matched Studies with Improved Internal and External Validity

Half Day | 1:00 pm – 5:00 pm

Franklin Hall 13, 4th Floor

José R. Zubizarreta, Harvard University

**Description:** In observational studies of causal effects, matching methods are extensively used to approximate the ideal study that would be conducted if controlled experimentation was possible. In this short course, we will explore recent advancements in matching methods to design matched studies with improved internal

and external validity. With these matching methods, we will: (1) directly obtain flexible forms of covariate balance, ranging from mean balance to balance of entire joint distributions, (2) produce self-weighting matched samples that are representative of target populations by design, and (3) handle multiple treatment doses without resorting to a generalization of the propensity score, instead balancing the original covariates. We will discuss extensions to matching with instrumental variables, in discontinuity designs, and for matching before randomization in experiments. The methods discussed build upon recent advancements in computation and optimization for large data sets. We will use the statistical software package 'designmatch' for R.

Participants will gain a clear picture of role of matching for causal inferences, and its pros and cons. They will learn how to construct balanced and representative matched samples, improving on traditional matching methods on the estimated propensity score. The target audience of the workshop is applied researchers with quantitative training and familiarity with traditional regression methods. Facility with R is ideal, but not strictly necessary as well-documented step-by-step code will be provided.

### SC7. Smart Simulations with SAS and R

Half Day | 1:00 pm – 5:00 pm

Franklin Hall 4, 4th Floor

Mehmet Kocak, University of Tennessee Health Science Center

**Description:** In statistical methodology research and practice, simulations are among the ways to show operating characteristics of the proposed method against the existing methods or alternative approaches. Depending on the response variables of interest in such simulations, univariate or multivariate, iterative or non-iterative, simulation designs must be considered very carefully to produce generalizable and reproducible conclusions regardless of the simulation platform, and this task is much more difficult and under-recognized than typically thought. In this short course, we will introduce simple to more complex simulation designs and the importance of simulation size; we will describe potential pitfalls that may not be easily recognizable and suggest what metadata to be captured for a clear description of the simulation process and results. We plan to carry out examples both in SAS and R to show similarities and differences between the two platforms. In doing so, we will utilize Graphical Analytics techniques, which are indispensable components of statistical learning and practice, and must be made part of any simulation plans as well.

Course participants are highly encouraged to have a personal computer with at least one of SAS or R (and R-studio) installed to practice alongside the instructor as the following modules are being presented:

Module-1: Simulating data for univariate random variables following Gaussian Distribution, Student-t-Distribution, Gamma Distribution and its special cases, Beta Distribution, Binomial Distribution, Poisson Distribution, etc.

Module-2: Simulation designs for one-sample hypothesis testing for continuous, binary, and survival endpoints. In this module, we will also illustrate iterative simulation designs such as Phase-I Dose Escalation Design, and Simon's Two-stage designs.

Module-3: Simulation designs for two- or more-sample hypothesis testing for continuous, binary, and survival endpoints. One of the main focus here will be Empirical Power calculations for Randomized Clinical Trials.

Module-4: Simulation designs for Multivariate random variables and designs that require iterative processing. We will compare and contrast SAS and R as two simulation platforms and discuss ways to improve efficiency in simulation design and conduct.

# TUTORIALS

## T1. An Introduction to Causal Effect Estimation with Examples Using SAS® Software

Monday, March 25 | 8:30 am – 10:15 am  
Salon A, 5th Floor  
Michael Lamm, SAS Institute Inc.

**Description:** How can you estimate a causal effect from nonrandomized data? As statisticians and data scientists are increasingly tasked with analyzing data that come from observational studies rather than randomized experiments, this is a question of increasing importance. This tutorial provides an overview of methods for estimating causal effects for dichotomous treatments. In particular, it illustrates causal effect estimation by propensity-score-based matching, inverse probability weighting, and doubly robust methods by using examples relevant to the biological and life sciences. The analyses are performed using the PSMATCH and CAUSALTRT procedures in SAS/STAT® software. Also briefly discussed are approaches for constructing and evaluating the underlying models, comparisons of the estimation methods, and the assumptions required for identifying and estimating treatment effects.

## T2. Building Effective Data Visualizations with ggplot2

Monday, March 25 | 10:30 am – 12:15 pm  
Salon A, 5th Floor  
Lucy D'Agostino McGowan, Johns Hopkins Bloomberg School of Public Health

**Description:** “If you’re navigating a dense information jungle, coming across a beautiful graphic or a lovely data visualization, it’s a relief. It’s like coming across a clearing in the jungle.” – David McCandless.

The ability to create polished, factual, and easily-understood data visualizations is a crucial skill for the modern statistician. Visualizations aid with all steps of the data analysis pipeline, from exploratory data analysis to effectively communicating results to a broad audience. This tutorial will first cover best practices in data visualization. We will then dive into a hands on experience building intuitive and elegant graphics using R with the ggplot2 package, a system for creating visualizations based on The Grammar of Graphics.

## T3. Meta-Analysis of Clinical Trials: Effects-at-Random or Studies-at-Random?

Monday, March 25 | 1:45 pm – 3:30 pm  
Salon A, 5th Floor  
Jonathan J. Shuster, University of Florida

**Description:** Meta-Analysis and Systematic Reviews stand at the top of most “Evidence Pyramids”. Virtually all random-effects meta-analyses ever done (classical or Bayes) use the “Effects-at-Random” premise, where the random effect size for each study is drawn from an urn and the population mean of the urn is estimated. The almost never used “Studies-at-Random” instead presumes that the observed studies are a random sample of studies, drawn from a large conceptual urn of studies. The important distinction is that in the “effects-at-random” presumption, there can be no association between the random effect sizes and the study design parameters, which determine study weights. It is impossible to prove beyond a reasonable doubt that no such association exists. The framework for inference in studies-at-random, which estimates the mean outcome in the urn of studies, using the study sample sizes as its weights, offers many advantages over effects-at-random. We cite three here. First, in the target population, the mean of each completed study is known without error. Single-stage cluster sampling methods can easily be applied. Second, studies-at-random, but not effects-at-random, recognize that the study sample sizes are random variables, a source of variation conveniently not considered in effects-at-random. Third, the asymptotic distribution of effects-at-random, but not studies-at-random require either normal assumptions or large samples within studies. Both approaches are asymptotic in the number of studies being combined. Of note, we shall present two eye-opening real situations for effects-at-random, where keeping the point estimates as they were, but cutting the standard errors uniformly in half, cause a highly significant result to become non-significant. This cannot happen to studies-at-random. We shall apply studies-at-random methods to three situations: (1) Low event-rate binomial trials, (2) Trials with quantitative endpoints, and (3) Bland-Altman analysis with repeated measures within subjects. Unlike the classical repeated measures Bland-Altman methods, it is completely robust to the lack of independence within subjects.

## T4. Modern Multiple Imputation

Monday, March 25 | 3:45 pm – 5:30 pm  
Salon A, 5th Floor  
Michael R. Elliott, University of Michigan

**Description:** In the four decades since it was first proposed, multiple imputation has come to offer a comprehensive and practical solution to the problem of making statistical inference when missing data is present. This tutorial will provide a brief overview of the theoretical background behind multiple imputation, and then discuss a variety of practical implementations beyond the fully model-based setting, including use of chained equations, and predictive mean matching. We will conclude with a review of relevant software packages for creating and analyzing multiply imputed datasets, including SAS, R, and IVEWare.

## T5. A Primer on Python for Statistical Programming and Data Science

Tuesday, March 26 | 1:45 pm – 3:30 pm  
Salon A, 5th Floor  
Christopher Fonnesebeck, Vanderbilt University Medical Center

**Description:** Though Python is ostensibly a general-purpose programming language, it has quickly become a dominant language for machine learning and data science applications. This is due in part to its fundamental strengths as a high-level language, and in part to the powerful set of third-party packages that comprise the Python “scientific stack”. In this hands-on tutorial, we will first cover the fundamentals of Python programming, including data structures, control flow, functions, and classes, with particular attention paid to aspects of the language that is idiomatic. The second part of the course will comprise a survey of Python libraries that are relevant for modern data analysis, particularly in the context of data science and probabilistic programming. These include: NumPy, SciPy, Jupyter, pandas, dask, scikit-learn, PyMC3, matplotlib, Seaborn, and TensorFlow. Demonstrations will be motivated with real-data examples, using Jupyter notebooks to allow for interaction and experimentation.

## T6. Analysis of Patient-Reported Outcomes

Tuesday, March 26 | 3:45 pm – 5:30 pm  
Salon A, 5th Floor  
Joseph Cappelleri, Pfizer Inc  
Andrew G. Bushmakina, Pfizer Inc

**Description:** Patient-reported outcomes are often relevant in studying a variety of diseases and outcomes that cannot be assessed adequately without a patient’s evaluation and whose key questions require patient’s input on the impact of a disease or a treatment. To be useful to patients, researchers and decision makers, a patient-reported outcome (PRO) must undergo a validation process to support that it measures what it is intended to measure accurately and reliably. In this tutorial, the core topics of validity and reliability of a PRO measure will be discussed. In addition, the specialized topics of clinically important responder and clinically important difference on a PRO measure will be featured. Other analytic areas such as longitudinal data analysis of a PRO measure will be highlighted. Illustrations will be provided through real-life and simulated examples, including simulation-based learning of the methodologies. Material is drawn in part from the book “Patient-Reported Outcomes: Measurement, Implementation and Interpretation” (Cappelleri, Zou, Bushmakina et al. 2013).

# ROUNDTABLES

Roundtables Monday, March 25, 2019 | 12:15 pm – 1:30 pm  
Salon F, 5th Floor

## R1. From Working Group to Center: How to Establish and Grow Research Groups

Jason Roy, Rutgers School of Public Health

**Description:** Research working groups can be great breeding grounds for new ideas and cross-disciplinary collaborations. However, there are challenges in maintaining active participation, given competing demands on people's time. In this roundtable, we will discuss this in the context of our experience growing a causal inference working group and later establishing the Center for Causal Inference. Whether your goal is to have a small working group or to form a research center, you should come away with useful tips to increase participation and productivity.

## R2. Time Management: Taming Your Inbox

Elizabeth Stuart, Johns Hopkins Bloomberg School of Public Health

**Description:** Do you have trouble keeping email under control? Do things slip through the cracks? Do you worry about not being able to find time for focused and intense work ("important but not urgent"), due to the "urgent but not important" tasks that come up? This roundtable will discuss strategies for time management, especially for dealing with email and the tasks that come with it. Come with your strategies and hear from others about what works for them!

## R3. Practical Issues in Clinical Trial Design and Analysis for Precision Medicine

Peter F. Thall, M.D. Anderson Cancer Center

**Description:** This round table discussion will focus on methods for dealing with practical considerations that arise in the process of clinical trial design, conduct, and analysis, with particular attention to newer phase I-II and randomized trial designs that include subgroup-specific decision making. Depending on the attendees' interests, we will discuss a variety of recent developments, including designs discussed in the 2016 book 'Bayesian Designs for Phase I-II Clinical Trials' by Yuan, Nguyen and Thall.

## R4. Statistical Leadership

Bhramar Mukherjee, University of Michigan Rogel Cancer Center

**Description:** In this discussion, I will focus on two types of leadership positions, an outward leadership challenge as a statistician leading a group of non-quantitative but exceptionally talented biomedical researchers and an inward leadership role as a Department chair to lead a Biostatistics department to the next generation. In the era of health big data, we need to lead as both an active doer and a careful thinker. I have benefitted greatly in my own career by simply being at the table where scientific strategies are being defined and scientific discoveries are being made. As a statistician, it is important to be in the room where it happens and truly learn to embrace, adopt and appreciate the diversity in data science and in the society!

## R5. Strategies for Success in Publishing

Heping Zhang, Yale University

**Description:** Publishing in peer-reviewed journals is a fundamental expectation for all biostatisticians. Understanding how peer-review and the editorial process works, and what makes an effective journal articles critical for success in publishing. This roundtable will focus on a range of strategies for successful publication of your research.

## R6. Submitting your Grant to NIH

Peter Kozel, NIH/NIDDK

**Description:** Have you ever been confused by the NIH grant system? Want to know tips for working out the appropriate institute and funding opportunities to submit to? Interested in what happens after you submit your application? The objective of this roundtable is to raise an awareness of how the NIH peer review process works, and to discuss some general do's and don'ts of application submission.

## R7. Analytic Challenges of Administrative Health Data

Rebecca Hubbard, University of Pennsylvania

**Description:** There is currently enormous enthusiasm for conducting research using data from electronic health records (EHR). However, analyzing EHR data entails many practical challenges. This roundtable will discuss key challenges for the analysis of EHR data including: missing and mismeasured variables, confounding by indication, and informative observation processes. The objective of the roundtable will be to raise awareness about concerns arising in the analysis of EHR data and to share best practices for addressing these challenges.

## R8. Tips for Interviewing Well

Joseph C. Cappelleri, Pfizer Inc

**Description:** This roundtable will focus on ways to interview well and therefore increase the chance of receiving a job offer. Many tips will be provided and discussed. Among them are researching the industry and the institution, clarifying your "selling points" and the reasons you want the job, and preparing for common interview questions. Basic and subtle interviewing skills will be discussed including (among others) how to make a good first impression, to get on the same side as the interviewer, and to empower yourself through thinking positive.

## R9. Effective Collaboration and Statistical Leadership—in Drug Development and Beyond

Lei Shen, Eli Lilly and Company

**Description:** Modern day drug development is highly complex and requires deep technical expertise in many scientific disciplines as well as effective collaboration among teammates with different expertise. The skills to address challenges that inevitably come up along the way are not the monopoly of any single discipline, but a case can be made that statisticians - thanks to our training in uncertainty quantification and ability to think in probabilistic terms - are in a prime position to step up as problem solvers. Somewhat paradoxically, our statistical leadership can be greatly enhanced by not being merely "statisticians who work in drug development", but rather "drug developers who happen to be statisticians". This roundtable gives us the opportunity to discuss and debate the requisite skillsets for us to develop - and conscious effort to make - to be successful in this realm. And are there close parallels in other industries, in government agencies, and in academic research?



# CENS MISSION



CENS seeks to advocate for the needs and concerns of students and recent graduates in collaboration with ENAR's Regional Advisory Board. Through annual events at the ENAR Spring Meeting, CENS strives to promote the benefits of participating in the ENAR community, support the advancement of students and recent graduates, and facilitate stronger connections within the statistical community.

## CENS Events at ENAR 2019

**CENS Sponsored Session:** Tuesday, March 26th from 3:45-5:30 PM.

Developing Collaborative Skills for Successful Careers in Biostatistics and Data Science Collaborations are an essential element of a productive career, especially for biostatisticians and data scientists. For emerging and new biostatisticians/data scientists, it can be difficult to develop, maintain, and benefit from good collaborations. It can be equally as difficult to recognize, navigate, and end poor collaborations. In this session, biostatisticians and data scientists from various career paths will draw from their own experiences to offer advice on collaborative skills necessary for a successful career.

**Networking Mixer:** Monday, March 25th from 5:30-6:30 PM.

All students and recent graduates are invited to attend the CENS Networking Mixer. Registration is not required - so please plan to attend!

**Networking Lunch:** Tuesday, March 26, 2019 from 12:30 PM - 1:30 PM at local restaurants.

CENS will organize lunches for groups of attendees that share similar interests. The goal is to help attendees meet and network with each other. Although CENS will help to coordinate lunch at local restaurants, please note that lunch is at your own expense and CENS will not be able to cater to special dietary requirements. Closer to the meeting time, CENS will email all attendees interested in this networking event

to request information to set up the groups and the lunch reservations. Participants meet at the CENS table in the Exhibition Hall at 12:15 PM before walking with their assigned group to a nearby restaurant for networking and lunch! Participation is open to all meeting attendees. If you would like to participate, please select the CENS lunch option on the registration form or email CENS at [enar.cens@gmail.com](mailto:enar.cens@gmail.com).

## About CENS

CENS was formed in 2012 by ENAR's Regional Advisory Board (RAB) to help ENAR better address the needs of students and recent graduates. CENS is composed of 10 graduate students, post-doctoral fellows, or recent graduates who are ENAR members. With the help of the RAB Liaison, CENS members collaborate to bring student/recent graduate concerns to the attention of RAB and ENAR; work to help ENAR better serve all students/recent graduates; advise and help implement ideas to enhance the benefits of ENAR membership and to increase awareness of the benefits of ENAR membership to students; organize a CENS sponsored session at each ENAR Spring Meeting; assist in planning events that help advance students' and recent graduates' education and careers; and contribute to the development of ENAR's social media presence.

## Join CENS

We are actively recruiting new members! Each member is appointed to a 2-year term. Within CENS, three or four people are chosen to participate in the steering committee, which reports to the RAB chair. Members of the steering committee will serve an additional year on CENS. CENS members meet in person yearly at the ENAR Spring Meeting and participate in conference calls throughout the year to plan events and address issues as they arise. If you are interested in joining CENS, please email [enar.cens@gmail.com](mailto:enar.cens@gmail.com).





# ENAR 2019 SPRING MEETING

WITH IMS & SECTIONS OF ASA

MARCH 24-27, 2019

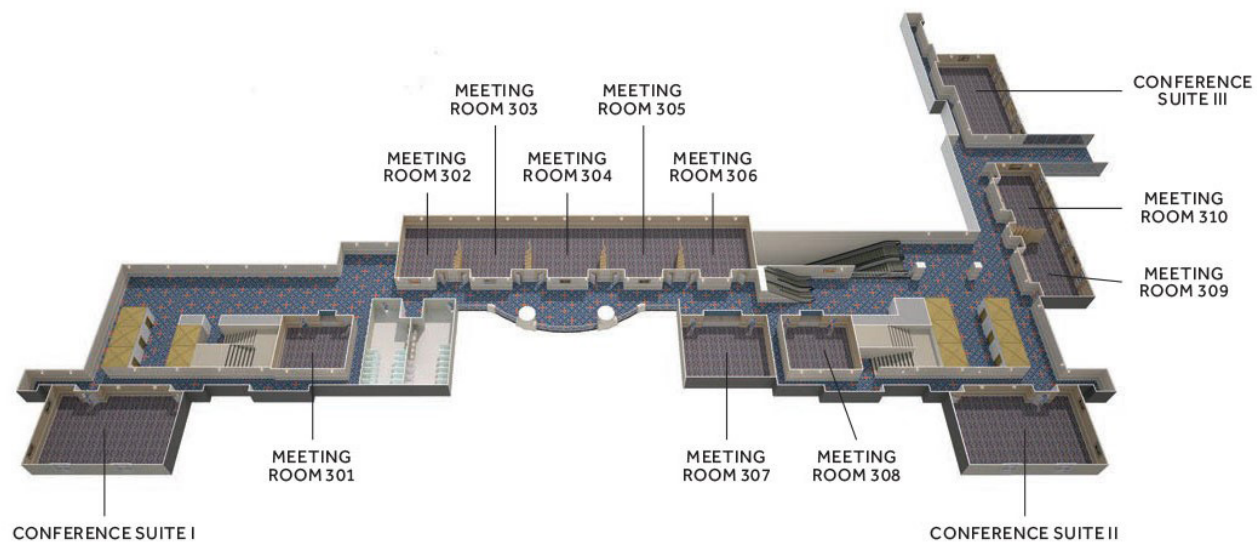
MARRIOTT PHILADELPHIA, PHILADELPHIA, PA



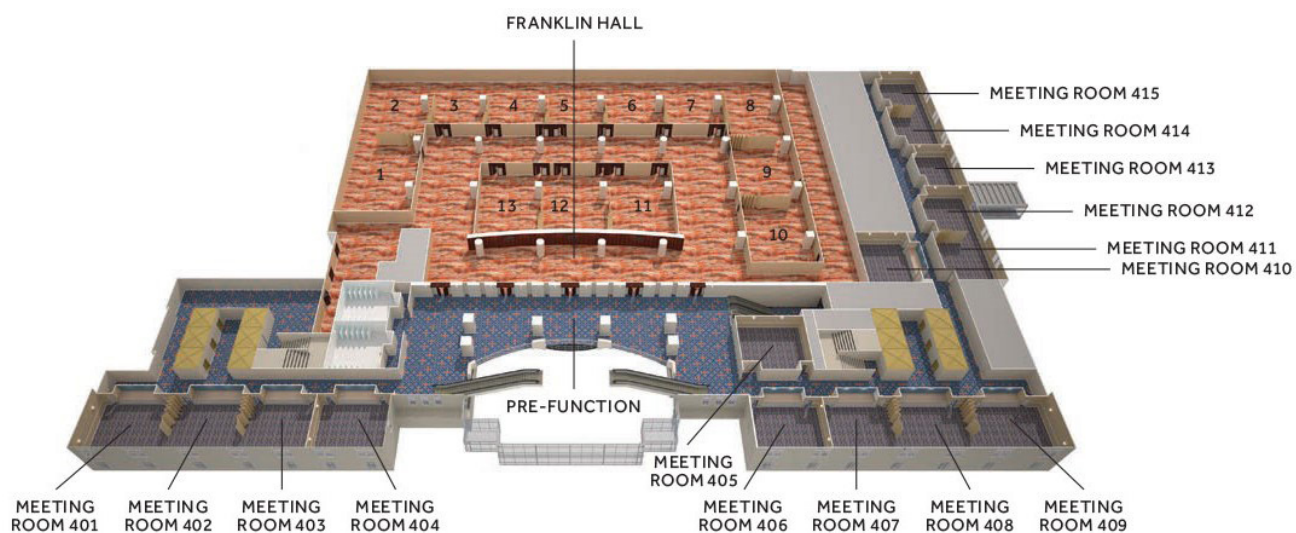


# FLOOR PLAN

## LEVEL 3

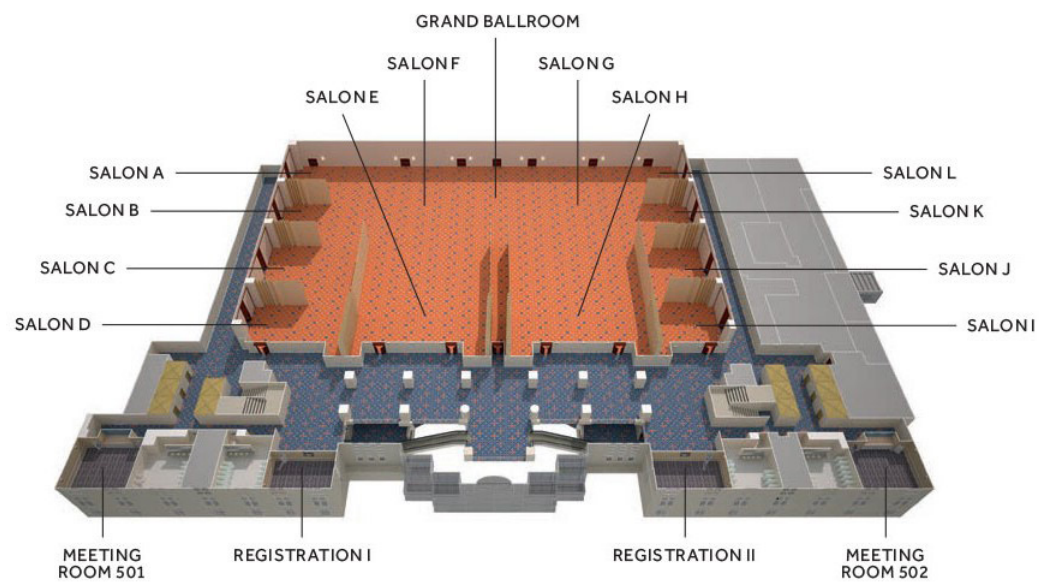


## LEVEL 4



# FLOOR PLAN

## LEVEL 5







11130 SUNRISE VALLEY DRIVE  
SUITE 350  
RESTON, VIRGINIA 20191  
PHONE: 703.437.4377  
FAX: 703.435.4390  
WWW.ENAR.ORG

