The National Academies of
SCIENCES • ENGINEERING • MEDICINE

Committee on National Statistics, Committee on Population, and the Board on Behavioral, Cognitive, and Sensory Sciences
Division of Behavioral and Social Sciences and Education

Agenda

Social Science Modeling for Big Data in the World of Machine Learning
for the National Institute of Aging
October 24–25, 2019
National Academy of Sciences Building
2101 Constitution Avenue NW, Washington, DC

Day 1: Thursday, October 24, 2019
Room 125

9:30am **Welcome, Introductions, and Workshop Goals**
Brian Harris Kojetin, *Committee on National Statistics*
Dana Plude, Partha Bhattacharyya, and Jonathan King, *National Institute of Aging*
Harlan M. Krumholz, *Yale School of Medicine (planning committee co-chair)*

10:00 **Setting the Rules of the Road—Avoiding Methodological and Data Errors While Implementing Machine Learning Techniques on Large Datasets**
Sherri Rose, *Harvard Medical School (planning committee member)*
Ziad Obermeyer, *UC Berkeley School of Public Health (planning committee member)*

11:15 *Networking Break*

11:30 **Standards for Healthcare Data Interoperability**
Josh Mandel, *Microsoft Healthcare [presenting remotely]*

12:00 **Navigating Data Access and Acquisition: Building Effective Data Partnerships**
Katy Haynes, *Nightingale*

12:30pm Lunch Break, *Food is available for purchase in the Refectory (lower level), or food trucks on 21st St and Virginia Ave NW*

1:30 **Case Studies I—Data Infrastructure Needs in Harnessing Social, Online, and News Data for Public Health Research**
Elaine Nsoesie, *Boston University [presenting remotely]*
Munmun De Choudhury, *Georgia Tech*

2:45 *Networking Break*

3:00 **Making Data Analysis Interpretable with Causality and Explanations**
Sudeepa Roy, *Duke University (planning committee member)*

3:45 **Case Studies II—Machine Learning Techniques for Medical Diagnoses and Treatment**
Bobak Mortzavi, *Texas A&M University*
Emma Pierson, *Stanford University*

5:00 **Sponsor reactions to Day 1; General Q&A**

5:15pm ***Adjournment***

Day 2: Friday, October 25, 2019
Room 125

9:00am **Brief Introduction to Day 2**
Sendhil Mullainathan, *University of Chicago (planning committee co-chair)*

9:15 **Recent Advances in Statistical Methodology for Machine Learning with Big Data**
Rajesh Ranganath, *New York University*

10:00 **Machine Learning for Regulatory Science: Assessing the Quality of Hospital Care**
Sharon-Lise Normand, *Harvard Medical School [presenting remotely]*

10:45 *Networking Break*

11:00 **Weak Supervision in Medicine: Optimism and a Pitfall**
Christopher Ré, *Stanford University [presenting remotely]*

11:45 **Developing Digital Measures from Person-Generated Health Data**
Luca Foschini, *Evidation Health*

12:30pm Lunch Break, *Food is available for purchase in the Refectory (lower level), or food trucks on 21st St and Virginia Ave NW*

1:30 **Roundtable Discussion: Lessons Learned and the Way Forward for NIA**
Harlan M. Krumholz, and Sendhil Mullainathan, *moderators*

3:00 ***Adjournment***

WORKSHOP PLANNING COMMITTEE

Harlan M. Krumholz (co-chair), Harold H. Hines, Jr. Professor of Medicine and Epidemiology and Public Health, Yale University School of Medicine

Sendhil Mullainathan (co-chair), Roman Family University Professor of Computation and Behavioral Science**,** University of Chicago Booth School of Business

Hedwig (Hedy) Lee, Professor of Sociology**,** Washington University in St. Louis

Ziad Obermeyer, Acting Associate Professor of Health Policy and Management, University of California, Berkeley School of Public Health

Sherri Rose, Associate Professor of Health Care Policy**,** Department of Health Care Policy, Harvard Medical School

Sudeepa Roy, Assistant Professor, Department of Computer Science**,** Duke University