Xu Shi, PhD

Assistant Professor, Department of Biostatistics at University of Michigan

February 6, Thursday 9-10 am EST

For more information regarding upcoming webinar schedule, please contact medecc.us@boehringer-ingelheim.com

Title

Beyond Curve Fitting: Transfer Learning and Causal Reasoning

Abstract

The growing availability of administrative healthcare data such as the electronic health record (EHR) data is opening new opportunities for research. EHRs are routinely collected longitudinal data containing demographics, medical diagnosis and procedure, medication, immunization, laboratory test results, radiology images, vital signs, and billing information. Ongoing efforts have been made to integrate large scale EHR data across healthcare systems, as well as to link EHR data with biobank, insurance claims, registries, and death indices. With such cost-effective data sources, the health of an individual is now characterized with unprecedented precision and depth, facilitating contemporary research that makes the transition from data to knowledge. However, EHR data are not collected for research purposes. Critical issues such as data quality, system heterogeneity, unmeasured and mismeasured confounding, high-dimensional covariates, and patient privacy concerns naturally arise. In this talk, I will detail the problem of inconsistent "languages" used by different healthcare systems, which limits the transportability of phenotyping algorithm and statistical method across different systems. I will present an automated data quality control and harmonization pipeline that aim to overcome this challenge. I will also present a tailored causal inference method that leverages the unique pool of information in EHR data to mitigate unmeasured confounding.

Professional Biography

Xu Shi is an Assistant Professor in the Department of Biostatistics at University of Michigan. She received her PhD in Biostatistics from University of Washington, Seattle in 2017. Prior to joining the faculty, she was a postdoctoral fellow at the Harvard Data Science Initiative at Harvard University. Her research focuses on developing novel statistical methods that provide insights from high volume and high variability administrative healthcare data such as the electronic health records (EHR) data. She is particularly interested in developing causal inference methods tailored to EHR data, automated knowledge extraction, data harmonization across healthcare systems, post-marketing drug safety surveillance, and high-throughput comparison of healthcare utilization. She is the Advanced Analytics Core Co-Lead of the FDA's Sentinel Initiative Innovation Center, developing innovative statistical methods to monitor the safety of FDA-regulated medical products and exploring novel ways to utilize information from EHR data.

Sponsored by

- American Statistical Association (Boston, Connecticut, Florida, New Jersey, Princeton/Trenton, and Washington chapters)
- Boehringer Ingelheim Pharmaceuticals, Inc. (Biostatistics and Data Sciences Department)
- New England Statistical Society (NESS)

For interested participants

Please send your email message to medecc.us@boehringer-ingelheim.com with the subject line "I want to attend no.21 on February 6th, please send me the link".

Boehringer Ingelheim

ASA AMERICAN STATISTICAL ASSOCIATION Promoting the Practice and Profession of Statistics®

NESS | New England Statistical Society