

## Great Plains IDeA-CTR Monthly Lecture

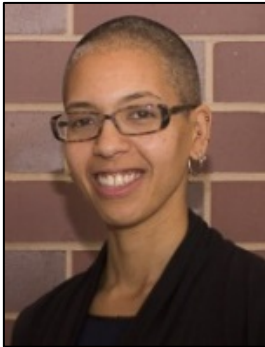
### The Role of Data Provenance in the Estimation and Analysis of EHR-Derived Phenotypes

Friday, April 26<sup>th</sup>, 2019 | 3:00-4:00 PM

Maurer College of Public Health (MCPH) 3013

Or, join via Zoom: <https://unmc.zoom.us/j/4025522260>

University of Nebraska Medical Center



**Rebecca Hubbard, PhD**

Associate Professor of Biostatistics  
Department of Biostatistics and Epidemiology  
University of Pennsylvania School of Medicine

Dr. Hubbard's research focuses on development and application of statistical methodology for studies that use observational data from clinical medical practice. This work encompasses evaluation of screening and diagnostic test performance, methods for comparative-effectiveness studies, and health-services research. A wealth of data on healthcare utilization, performance and outcomes are available from electronic health records. However, the complex observation scheme that gives rise to these data necessitates careful study design and analysis to ensure that inference is valid. Dr. Hubbard's methodological research emphasizes development of statistical tools for such inference and has been applied to studies of cancer screening, aging and dementia, pharmacoepidemiology, women's health and behavioral health.

**This talk will:**

- Introduce data provenance, the process that gives rise to observed data, is complex for electronic health record (EHR)-derived data and can lead to patterns of missing data that must be accounted for in analyses.
- Discuss alternative methods for deriving phenotypes from EHR, highlighting the role of data provenance in each.
- Illustrate the implications for bias and hypothesis testing of using imperfect EHR-derived phenotypes.
- Demonstrate a straightforward approach to correct for phenotyping error when using EHR-derived phenotypes.

Trouble viewing? Call 402-559-8090

Great Plains IDeA-CTR Network  
<https://qpctr.unmc.edu/> | 402.552.2260