

**Department of Epidemiology and Biostatistics
Biostatistics Seminar**

Wednesday, September 23, 2015
12:00pm - 1:00pm -- WG73

“Large-scale extraction of drug side effects from heterogeneous data resources and its implication in drug discovery”

Rong Xu, PhD. Assistant Professor
Department of Epidemiology and Biostatistics
Case Western Reserve University

Abstract: Systems approaches to studying drug phenotypes, specifically drug-side-effect (drug-SE) associations, are emerging as an active research area for drug target discovery, drug repositioning, and drug toxicity prediction. However, currently available drug-SE association databases are far from being complete. In this talk, I will present natural language processing techniques (NLP), text mining, data mining, and machine learning techniques that we recently developed to extract drug-SE associations (2,905,811 drug-drug-SE pairs representing more than 6,000 drugs) from three large-scale, independent, and complementary data sources, including 21 million published biomedical literature, 44,979 FDA drug labels, and 4.3 million patient records from U.S post-marketing drug safety surveillance system. I will demonstrate how the drug-SE association data, in combination with other drug- and disease-related data, can be used for drug discovery.