Breakout Session I

*Big Data Management for Consortia*

Uri Kartoun, PhD and Andrew T. Chan, MD

Friday, October 20, 2017

11:00 AM – Noon
In this session, two leading investigators will discuss access, design and management of large data sets derived from consortia and electronic health records (EHR) for application to research in liver diseases. Dr. Andrew Chan has over a decade of experience working with “big” data sets derived from large, prospective longitudinal cohort studies enrolling several hundred thousand individuals. These cohorts include the Nurses’ Health Studies I and II, the Health Professionals Follow-up Study, the Physicians’ Health Study, the Prostate Lung Colorectal Ovarian Trial, and the Framingham Heart Study. He is also a study chair and investigator within three large consortia, the Genetic Epidemiology of Colorectal Cancer (GECCO), International Survival Analysis of Colorectal Cancer (ISAAC), and the National Cancer Institute Cohort Consortium. Dr. Chan will discuss the history and maintenance of these cohorts. He will also provide insight into how to access and manage data from these cohorts for the development of individual research studies.

Dr. Kartoun is a research staff member at IBM Research in Cambridge, MA. Previously he was a research fellow at Harvard Medical School/Massachusetts General Hospital. Dr. Kartoun will focus on the development of datasets leveraging information within EHR collected during usual clinical care. These data represent a powerful resource to study diseases and their associated comorbidities at population scale in a real-world context. Recent studies have demonstrated the utility of EHR analysis to assess associations between patient characteristics and outcomes, identify more precisely individuals at high risk for a disease, and confirm beneficial or harmful effects of medical treatments. Dr. Kartoun will discuss the design, development and maintenance of EHR databases, including the specific issues regarding handling of sensitive personal information, challenges of processing large collections of records in a timely manner, and application of relevant computational methods. This will include a discussion of approaches such
as machine learning and text processing algorithms to access data from narrative clinical notes to achieve higher accuracy of diagnosis and prediction of outcome. Together, Dr. Chan and Dr. Kartoun will compare and contrast the strengths and weaknesses of various databases to conduct research in liver disease.
Breakout Session II

Animal Models and Translation to Human Liver Disease

Bryan C. Fuchs, PhD and Anna Mae Diehl, MD, FAASTLD

Friday, October 20, 2017
11:00 AM – Noon
Overview of Animal Models Available

- Diet-based - high fat diet, choline deficient diets (MCD, CDAA, CDAHFD), western diets, sucrose and fructose, cholesterol
- Genetic – ob/ob, PTEN

How should we determine which models best reflect human disease?

- Transcriptomic analysis
  - Very little overlap between animal models and human tissues (Teufel et al. Gastroenterology 2016)
- Histologic analysis
  - While Western diet better mimics metabolic features of the disease, MCD diet produces a more reliable NASH phenotype (Machado et al. PLoS One 2015)
- Is it even possible to generate a model with reliable and robust metabolic and histologic features of the disease?

Which models should we therefore choose for testing treatments and how should we diagnose and monitor treatments in these models?

- Should we just rely on histologic stains and traditional RNA biomarkers like Col1a1 and Acta2?
- Should we evaluate clinical readouts like MRE, PDFF, serum biomarkers?
- What about molecular MR Imaging to detect collagen (Fuchs et al. J Hepatol 2013) or myeloperoxidase (Pulli et al. Radiology 2017)?