BIG DATA: CURRENT STATUS AND FUTURE DIRECTIONS AGENDA

Organizers: Michael F. Chiang, MD, Anne L. Coleman, MD, PhD, FARVO and Seth Blackshaw, PhD

Big Data is one of the most frequently used terms in the press. This course defines Big Data and describes Big Data sets derived from clinical data available in eye and vision research. How to visualize Big data and applications of machine learning are also discussed.

Introduction and Welcome
Anne L. Coleman, MD, PhD, FARVO, Jules Stein Eye Institute, UCLA

A general overview of Big Data, its definition, uses, benefits and liabilities will be presented.

Tapping into Health Care Claims Databases to Learn About Ocular Diseases
Joshua D. Stein, MS, MD, Kellogg Eye Center, University of Michigan

This talk will describe some of the common large health care claims datasets that researchers can access to study patients with ocular diseases. We will go over some of the strengths of performing research using these databases over other sources as well as limitations of working with these resources. Examples of analyses performed using health care claims data will be described so the audience can get a flavor for the types of studies and research questions that can be answered using these data sources.

Mining electronic health record (EHR) data
Michael F. Chiang, MD , Casey Eye Institute, Oregon Health & Science University

This presentation will discuss practical applications of mining EHR data. These will include clinical research applications of large-scale EHR data, as well as operational applications of EHR data for improving real-world clinical workflow.

Clinical Registries and "Big Data"
Paul P. Lee, JD, MD, FARVO, Kellogg Eye Center, University of Michigan

Clinical registries for eye care have existed since the post US Civil War period's trauma registry. In the last 40 years, several eye registries have been developed, covering not only trauma and ocular medications but also focused clinical issues such as cataract and cataract surgery. The Swedish National Cataract Registry is one excellent example. The newest registries, such as the American Academy of Ophthalmology's IRIS and the American Optometric Association's MORE are even more comprehensive, expanding registry coverage to the broad range of eye care and eye diseases that affect human health. The advent and growth of these comprehensive registries will enable us to build upon the quality and safety work that has been done for cataracts and ocular medications to the broad range of eye conditions as well as to begin to explore detailed relationships among many factors that have heretofore been difficult to understand due to practical limitations such as sample size and the necessity of chart abstraction. As such, clinical registries are both a powerful tool and exemplar of the promise of "big data" to improve vision and lives.

Visualizing "Big data"
Aaron Y. Lee, MD, University of Washington

One of the great challenges in Big Data is being able to visualize large volumes of data in order to understand new insights and trends. This presentation will feature several visualizations aiming to help researchers slice through different dimensions of data. Many of these visualizations are interactive allowing for deeper hypothesis generation by being able to visualize data.

Applications of machine learning and deep learning
Adnan Tufail, FRCOphth, MD, Moorfields Eye Hospital

The current explosion of activity in machine learning is remarkable. A non-technical overview of machine learning aimed at clinicians will be given and the potential to capturing the full value of big data discussed. We will discuss the importance of validation of machine learning algorithms. Current and emerging applications of machine learning in ophthalmology, such as the grading of diabetic retinopathy and diagnosis of AMD on retinal images, will be given.

Big data and ophthalmic genetics
Michael B. Gorin, MD, PhD, FARVO, Jules Stein Eye Institute, UCLA

There is considerable interest in multiple areas of medicine to use big data from electronic health records and genetic data. There are a number of opportunities and challenges that confront the use of this approach to address key issues in ophthalmic genetics and foster the application of molecular genetics in ophthalmic care. A number of projects have successfully used selective data from large data sets to test for associations with genetic variants/markers. We will review recent efforts in medicine and ophthalmology to employ institutional and consortia biobanks and EMR data in conjunction with genome wide, next generation sequencing and/or genotyping to elucidate genetic components of disease. This talk will highlight the challenges of defining phenotypes for disease using ICD-9, ICD-10 codes and natural language processing, harmonizing ocular-related measures that are not standardized and determined by multiple methods and devices, dealing with the laterality of diagnoses and treatments, missing and incomplete data especially with records that rely on varying inclusion of information from different medical specialties, and the lack of adequate family historical data. Realization of the potential of big data for ophthalmic genetics will ultimately require a concerted effort to address the consistent acquisition of data for clinical records.

Panel discussion - Brainstorming on next steps and global collaborations
Panelists: Anne L. Coleman, MD, PhD, FARVO, Joshua D. Stein, MS, MD, Paul P. Lee, JD, MD, FARVO, Aaron Y. Lee, MD, Adnan Tufail, FRCOphth, MD, Michael B. Gorin, MD, PhD, FARVO, Seth Blackshaw, PhD, David Cobrinik, MD, PhD, Salil Anil Lachke, PhD, Jiang Qian, PhD, Heng Zhu, PhD

Moderator: Michael F. Chiang, MD