Comprehensive Research Experience for Medical Students
Summer Research Program 2020

Supervisor/Project Information Form
Due February 6, 2020 by email to crems.programs@utoronto.ca

Supervisor Name: Geoffrey Liu (Oncology, Medicine) and Wei Xu (Biostatistics)

Project Title: Machine Learning and Biostatistics for Analysis of Symptoms and Toxicity (MBLAST) in Lung and Head and Neck Cancer

Hospital/Research Institution: University Health Network/Princess Margaret Cancer Centre

Email: geoffrey.liu@uhn.ca

Field of Research (2 keywords): cancer outcomes, data science

Department: Medicine, Epidemiology (Dalla Lana School of Public Health)

School of Graduate Studies Appointment (IMS, LMP, IHPME etc)? Yes If YES, please name: Medical Biophysics, Epidemiology, IMS, and Pharmacology and Toxicology

Brief Project Description (273 words): This innovative project intersects the use of clinical data (both structured and unstructured data) from the electronic health record of lung cancer patients with machine learning and biostatistical methods. Currently the use of Google Analytics after retail data and pattern-recognition (such as facial recognition/driver-less cars) has revolutionized these areas. We aim to apply similar methods to analyze symptoms and toxicities of Princess Margaret data on lung and head and neck cancer. We have built the data lake that contains our medication ordering data (OPIS program), laboratory data, MOSAIQ (radiation), patient-reported symptom and other clinical data (see 2BLAST catalyst project on http://pmdatascience.ca/projects/). The student will track the steps used by our data abstractors for structured (i.e. laboratory) and unstructured (i.e. free-text), and with our clinical teams to describe the algorithms for identifying symptoms and toxicity and attributions to specific therapies. The student will therefore bridge between the clinical teams with our data scientists who will be programming the first-level baseline algorithms that starts the machine learning process. The student will therefore be learning clinical research, oncology patient management, in combination with innovative machine learning methods. Co-Supervisors Dr. Liu and Xu have a longstanding history of supervising medical students locally, nationally, and internationally (see www.uhncombiel.com) with students presenting their work in abstracts and published manuscripts; 78% have continued academic research activity after CREMS. The student will have hands-on clinical experience that will inform the data abstraction process, with direct interaction with data scientists and programmers. This project is for a student who has strong inherent interest in oncology (medical, radiation, surgical), clinical research, programming and data science. Programming knowledge is not required, but helpful.