



Comprehensive Research Experience for Medical Students
Summer Research Program 2019

Supervisor/Project Information Form

Due February 20 2019 by email to crems.programs@utoronto.ca

Supervisor Name: Robert Wu

Project Title: Identification of patients admitted with COPD exacerbations and stratification of those at high risk of readmission using natural language processing and machine learning

Hospital/Research Institution: Toronto General Research Institute

Email: Robert.wu@uhn.ca

Field of Research (2 keywords): COPD, machine learning

Department: Medicine

School of Graduate Studies Appointment (IMS, LMP, IHPME etc)? Yes/No: Yes If YES, please name: IMS

Project Title: Identification of patients admitted with COPD exacerbations and stratification of those at high risk of readmission using natural language processing and machine learning

Brief Project Description (< 300 words):

Background: Patients with Chronic Obstructive Pulmonary Disease (COPD) who are admitted to hospital are at high risk of readmission. While therapies have improved and there are evidence-based guidelines to reduce readmissions, there are significant challenges to implementation including 1) identifying all patients with COPD early in admission to ensure evidence-based, high value care is provided and 2) identifying those who are at high risk of readmission in order to effectively target resources.

Objective: Using machine learning and natural language processing, we will develop models to 1) identify all patients with COPD exacerbations admitted to hospital and 2) stratify them to distinguish those who are at high risk of readmission

Methods: From Toronto hospitals, we will develop a very large dataset of patient admissions for all medical conditions including exacerbations of COPD from the electronic health record. This data will include both structured data such as age, gender, medications, laboratory values, co-morbidities as well as unstructured data such as discharge summaries and physician notes.

Using the dataset, we will train a model through natural language processing and machine learning to be able to identify people admitted with COPD exacerbation and identify those patients who will be at high risk of readmission within 30 days. We will test the ability of these models to determine our predictive accuracies.

Student's role: The student will be part of a team working with clinicians and machine learning experts. The student's responsibilities will include writing code to prepare the data and to conduct preliminary analyses using machine learning libraries. The student will gain a strong working knowledge of machine learning and natural language processing through this project.