Comprehensive Research Experience for Medical Students
Summer Research Program 2021
Supervisor/Project Information Form

Due February 24, 2021 by email to crems.programs@utoronto.ca

**Supervisor Name:** Cari Whyne (Co-Supervisor Elias Khalil, MIE)

**Project Title:** A Machine Learning Approach to Predict Healthcare Resource Utilization for Patients Undergoing Total Knee Arthroplasty

**Hospital/Research Institution:** Sunnybrook Research Institute

**Email:** cari.whyne@sunnybrook.ca

**Field of Research (2 keywords):** Orthopaedics, Machine Learning

**Department:** Surgery

**School of Graduate Studies Appointment (IMS, LMP, IHPME etc)?** Yes

**If YES, please name:** IMS
**Brief Project Description (<300 words):**

Background: Total Knee Arthroplasty (TKA) is one of the highest volumes and resource intensive surgical procedures. As such, it has been targeted by the Ministry of Health to advance its agenda of quality care and fiscal sustainability. Two important drivers of the cost of surgical care are operative time (OT) and post-operative inpatient length of stay (LOS). Therefore, it is critical to be able to predict TKA OT and LOS for optimal resource allocation.

Objectives: 1) Create a machine learning (ML) model that predicts TKA OT and LOS, 2) Identify and quantify the impact of key drivers contributing to OT and LOS on scheduling.

Methods: Preoperative demographic and clinical variables from all patients undergoing unilateral TKA between 2014-2018 in the National Surgical Quality Improvement (NSQIP) database will be used to predict OT and LOS. The dataset of over 200,000 procedures will be split across demographic parameters into a training, validation, and testing dataset with a ratio of 70%, 15%, and 15%, respectively. ML models that will be tested include logistic regression, random forest, and neural networks. The quality of the ML models will be analyzed using the area under the receiver operating characteristics on the validation dataset. The efficacy of the optimal model will be evaluated on the test dataset, considering accuracy, precision, sensitivity and specificity of the model. The optimal ML models to predict the OT and LOS of surgical procedures will be evaluated on local hospital data in order to simulate the scheduling of the operating room and bed capacity. The efficacy of the simulation by first predicting demand and then optimizing resources will be compared to the current state resource allocation system.

Student role: ethics approval, data analysis, and manuscript preparation. Supervisors will guide the student throughout the project with weekly meetings to ensure project success.