Replicability, Semi-Supervised Learning and Generative AI: Recent Statistical Work in Cancer Biostatistics

Naim Rashid, PhD

Associate Professor, Department of Biostatistics in UNC Gillings School of Global Public Health, with joint appointment at the Lineberger Comprehensive Cancer Center, and Grizzle Lecturer Awardee

In this talk we will discuss recent statistical work in several areas relevant to advancing cancer care and research. First, the replicability of statistical algorithms for clinical decision-making has been of significant recent concern in biomedical research, where multiple factors may limit the generalizability of models trained on individual studies. To improve replicability, we build upon our previous work to propose a scalable high dimensional data integration and multi-study learning approach for supervised learning problems with respect to a variety to survival and other clinical endpoints. We show that our method improves prediction performance in new studies, while also selecting features whose effects are consistently non-zero across training studies. We follow this discussion with an introduction to the use of semi-supervised and self-supervised learning to jointly identify lower dimensional features that are associated with clinical outcomes from genomic datasets, with applications to the identification of clinically relevant cancer subtypes. We end with a discussion on recent work in generative AI to empower cancer patients to find clinical trials that both meet their personal preferences and determine trial eligibility.

Thursday, April 125, 2024, 4:00-5:00PM Eastern

133 Rosenau Hall

Zoom Link: https://unc.zoom.us/j/98423779288?pwd=b0tqYThCQTAxeDdTQ0FRY3RnazdwQT09

Meeting ID: 984 2377 9288

Passcode: 631794

UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH Biostatistics