BSTA 750 Statistical Methods for Risk Prediction and Precision Medicine
Time: Fall (or Spring) 2021
Instructor: Jinbo Chen, Ph.D.

Course Objective

This is an advanced elective course for graduate students in Biostatistics, Statistics, Epidemiology, and other BGS disciplines. It will cover various topics for evaluating the performance of biomarkers to predict risk of clinical or disease outcomes, specifically including relative, absolute and competing risks for binary and time-to-disease outcomes; statistical inference for quantifying predictive accuracy with binary and time-to-event outcomes; statistical methods and inference for case-control study designs; Efficient study design issues for biomarker evaluation. This course is designed to help students 1) understand various concepts of risk in the medical literature; 2) understand various statistical methods for evaluating prediction performance of biomarkers and diagnostic tests and for designing efficient biomarker studies; 3) improve the ability to read critically papers published in statistical and medical journals on related topics; and 4) develop research ideas for risk prediction. Upon successfully completing this course, students will be able to: 1) Conduct statistical analysis for evaluating prediction performance of biomarkers and diagnostic tests; 2) Have a better ability to read and understand papers published in statistical and medical journals on related topics; and 3) Be well prepared to work on related topics for dissertation.

The course will be 0.5-unit and meet once or twice a week for a total of 3 hours for half a semester.

Pre-requisites: BSTA 630 and BSTA 632 or the equivalents; Or permission by the instructor;

Course topics

- *Risk measures/concepts that are commonly used in health research*
  - Relative risks, absolute risks, penetrance (genetics);
  - Competing risks
- *Commonly used study designs*
  - i) Cohort study
  - ii) Case-control and Cross-sectional studies
  - iii) Family-based study design (genetics)
- *Risk model evaluation*
  - Measures for evaluating the performance of risk prediction model and inferential methods
  - Statistical methods for ROC methodology for binary and time-to-event outcomes
  - Statistical methods and inference for case-control study designs
- *Efficient study designs for measuring predictive biomarkers*
  - i) Outcomes: Time-to-event outcomes:
    Designs: Case-cohort study design; stratified case-cohort study design;

Nested case-control study design

ii) Outcomes: Binary oucomes:
Designs: Two-phase cross-sectional or case-control study designs

iii) Study design and analysis issues for real-time risk predictions

- *Two Guest lectures on statistical methods for precision medicine*

Grading: Final grade will be determined as follows: 50% homework, 50% course project. There will no in-class midterm or final exam.

Homework: There will be three sequential homework assignments. They will involve a combination of analytical problems, data analyses, and simulation studies. The data analyses and simulation studies latter two align with the course project and are meant to track and help students' progress on the project. For problems involving programming, both the program and its output should be turned in. Students are encouraged to work together on homework; but each student must turn in his/her own solutions.

Course Project: Students are expected to work on a group project during the course and present the results as the final exam. Students will be divided into groups, each having 2-4 students depending on the total number of students. Each group chooses one topic that is covered in course lectures. The students are expected to identify a research topic by performing thorough literature review and perform simulation studies and real data analysis. The topic does not have to be novel, and preferably is related to students' own dissertation research or rotation projects. As the final exam, each group is expected to prepare a 15 minute presentation summarizing their main findings, and defend their work on both theoretical and practical aspects.

Textbooks: There is no textbook for this course. We will use class notes and heavily rely on journal articles.

***Scientific Rigor and Reproducibility*:** Through critical review of key literature and comprehensive data analysis of the course project, this course will provide rigorous training on various study designs and validation methods for developing accurate risk prediction model.