This course will present detailed discussion of the methods of predictive modeling, with applications to clinical and population health settings. Risk stratification models can be used to assess eligibility to clinical trials and interventions, to identify those at risk for more intensive therapy, and to guide prevention priorities. When does adding another variable to prediction improve classification? What are the trade-offs to consider when moving to implementation in real world settings?

The philosophy of the course is that learning should move away from the dictionary definition, “to receive instruction, to be informed, to commit to memory” to one that works for how computers learn, where committing to memory is a trivial task: things (students in a course like this?) learn when they change their behavior in a way that makes them perform better in the future. A performance based definition. We explore how prediction models fit into this concept of learning focusing on their development and implementation to improving health outcomes.

Building from traditional risk factor identification through regression analysis to model refinement and validation of prediction, a number of statistical approaches will be reviewed. Each method is motivated by clinical examples. Topics covered will include model development and validation; regression approaches, model selection, inference, averaging; classification and regression trees, and other topics. Approaches to validation will be discussed and strategies for estimation of added value with expanded variable lists will be a key focus of this applied quantitative methods course. Issues in moving to implementation and evaluation in clinical settings will round out the course.

Through examples, class discussion, and homework, students will become familiar with the methods for development, validation, and implementation of prediction models. Students will critically read and discuss a range of prediction model manuscripts. They will prepare an analysis plan for design and implementation of a prediction model to improve health outcomes.

**Course note:** Biostatistics I and II (M21-560 and M21-570) or equivalent are required prerequisites. SAS software required. (R or STATA could be used.)

**Competencies:** Develop the knowledge and skills to design, conduct, and analysis of risk prediction modeling, implement, and evaluate epidemiology-related, health services or clinical research projects of clinical or public health significance including:

- Obtain the knowledge and principles in basic issues involved in the clinical prediction rules, including design, development, validation, implementation, and interpretation of results for their application in clinical or public health settings.
- Develop the knowledge and skills with biostatistical methods and computer software for performing appropriate analyses of public health services or clinical outcomes data.
- Develop the knowledge and skills to critically evaluate existing risk prediction models and its translation to inform decision making of clinicians and policymakers.
- Apply principles of study design and evaluation to clinical research and implementation projects.
Text: (available on reserve at the library)


**Additional resources:** A more statistical theory based text for reference is

- Readings posted in each class

**Evaluation and Grading:** Grading will be based on 1) class participation (10%), 2) two homework assignments (20% each), and 3) a final project and presentation (50%).

**Class participation:** Focusing on applied learning, we will use published papers to promote class discussion. You must review each assigned paper in advance, be prepared to discuss the paper in class, and contribute to class discussion.

**Homework:** There will be two homework assignments, including data analysis and interpretation. Homework should be submitted by the due date indicated.

**Final project and presentation:** You will write a final paper that proposes a project to develop and/or validate a clinically relevant risk prediction model or update an existing model in your area of interest. The paper should provide rationale for developing, validating, or updating a risk prediction model; describe your method to develop and/or validate a model and alternative strategies assuming your original plan did not work; and include discussion on strengths and weaknesses of your method/model. The final paper should be 5 pages double spaced using Arial font size 11 with 1” margins and submitted by the due date as indicated. Rational/introduction should be less than 1 page. You will present your final paper to the class on an assigned date.

Mental Health Services are available for full-time students enrolled on the Medical School campus. Students can self-refer to a counselor (phone: 314-362-2404, Option # 1 or Option # 2); or make an appointment with Dr. Karen Winters through Student Health Services (SHS), telephone: 314-362-3523, and follow the prompts.

There are also contractual mental health service providers who are available off-campus. More information regarding this coverage and a list of participating providers are accessible via https://wusmhealth.wustl.edu/ and then clicking on Students and scrolling down to Mental Health Information https://wusmhealth.wustl.edu/students/mental-health-information/. Please do not hesitate to reach out to the MPHS program leadership (https://mphs.wustl.edu/contact/), or to any of our off-campus providers https://wusmhealth.wustl.edu/
Course Schedule (Tentative)

Week 1: Overview of Clinical and Public Health Risk Prediction

Reading:
- Chapter 2 in the Clinical Prediction Models (E.W. Steyerberg)
- Wald N. When can a risk factor be used as a worthwhile screening test? BMJ 1999; 319:1562-5

Week 2: Study Design and General Issues

Reading:
- Chapter 3, 4, & 9 in the Clinical Prediction Models (E.W. Steyerberg)

Additional reading:
- To be announced

Week 3: Risk Prediction Model Development and Evaluation

Reading:
- Chapter 7, 10, 11, 15.1 & 15.2 in the Clinical Prediction Models (E.W. Steyerberg)

Additional reading:
- To be announced

Week 4: Data Analysis Practice Session I – Model Development

GUSTO-I data

Reading:
- Chapter 22 in the Clinical Prediction Models (E.W. Steyerberg)

Week 5: Model Calibration and Validation

Reading:
- Chapter 15.3 & 17 in the Clinical Prediction Models (E.W. Steyerberg)

Additional reading:
Week 6: Class Discussions - Model Development and Validation

Reading:
- To be announced

Additional reading:
- To be announced

Week 7: Improving Existing Risk Prediction Models

Reading:

Additional reading:
- To be announced

Week 8: Data Analysis Practice Session II - Model validation

Reading:
- Chapter 22

Additional reading:
- To be announced

Week 9: Reporting a Prediction Model

Reading:
- Chapter 18 in the Clinical Prediction Models (E.W. Steyerberg)

Additional reading:
- To be announced

Week 10: Other Approaches for Prediction I: Classification and regression trees (Guest lecture)

Reading:

Additional reading:
- To be announced

**Week 11: Application of Risk Prediction Models in Clinical Settings** (Guest lecture)

Reading:
- To be announced

**Week 12: Your Disease Risk** (Guest lecture)

Your Disease Risk website: [https://siteman.wustl.edu/prevention/ydr/](https://siteman.wustl.edu/prevention/ydr/)

Reading:

**Week 13: Other Approaches for Prediction II** (Guest lecture)

Reading:
- To be announced

**Week 14: Risk Communication** (Guest lecture)

Reading:
- To be announced

**Week 15 & 16: Student Presentations**