

## The Certificate in Data Science in Health Economics and Outcomes Research

### About the Certificate

The Certificate in Data Science in HEOR is designed and taught by faculty from the CHOICE Institute, within the University of Washington's School of Pharmacy. It is offered through UW Professional & Continuing Education (UWPCE).

- The certificate comprises three online courses that are asynchronous and self-paced within the quarter dates:
  - Autumn quarter: September 24, 2025—December 12, 2025
  - Winter quarter: January 5, 2026—March 20, 2026
  - Spring quarter: March 30, 2026—June 12, 2026
- Each course includes two optional interactive virtual meetings with the instructor and teaching assistant.
- Courses include recorded lectures, readings, discussion activities, quizzes, and a quarterly individual capstone project.
- Students should expect to spend about 5-6 hours per week in course effort.
- All readings are available to read online or download from the Canvas course learning management system. There is no required textbook.
- Students have access to the University of Washington libraries.
- Courses are graded as Successful Completion (SC), Unsuccessful Completion (USC), or Incomplete (I).

### Courses (2024 curriculum content—subject to slight changes for 2025)

#### **HEOR DS 510:**

#### **Principles of Data Science and Health Care Data**

Autumn Quarter

Course 1 of 3 in the Data Science in HEOR Certificate Program

Instructor: Professor Beth Devine, PhD, PharmD, MBA, [bdevine@uw.edu](mailto:bdevine@uw.edu)

#### **Course Description**

This course provides an overview of the field of data science, laying a solid foundation for the material covered in quarters 2 and 3. The course includes an introduction to the broad field of data science; types of real-world evidence data sources; the details of data management; fundamental concepts of the scientific process, biostatistics and epidemiology; introduction to causal inference methods, which are covered in quarter 2; evidence synthesis; brief introduction

to artificial intelligence and machine learning and their application to healthcare, the material that will be covered in quarter 3.

## Course Learning Outcomes

By the end of the course, students will be able to:

- Compare and contrast sources of health data
- List the steps involved in data management
- Understand and explain fundamental scientific, statistical and epidemiologic concepts
- Interpret a network meta-analysis
- Appreciate the many potential applications of artificial intelligence and machine learning to the field of health economics and outcomes research

## Course Lessons

- Module 1 – Introduction to Data Science
  - Lesson 1, Part 1: Introduction to Data Science
  - Lesson 1, Part 2: Introduction to Data Science (continued)
- Module 2 – Data Sources
  - Lesson 2: Claims Data
  - Lesson 3, Part 1: Real-World Evidence
  - Lesson 3, Part 2: Real-World Evidence (continued)
  - Lesson 4: Electronic Health Record Data
- Module 3 – Data Management
  - Lesson 5, Part 1: Data Management
  - Lesson 5, Part 2: Data Management (continued)
- Module 4 – Fundamentals of the Scientific Process
  - Lesson 6: Fundamentals of the Scientific Process
  - Lesson 7: Introduction to Biostatistics
- Module 5 – Principles of Probability and Epidemiology
  - Lesson 8: Probability and Distributions
  - Lesson 9, Part 1: Principles of Epidemiology – Study Designs
  - Lesson 9, Part 2: Principles of Epidemiology – Bias, Confounding, Effect Modification
- Module 6 – Data Analysis
  - Lesson 10: Regression Methods
  - Lesson 11: Advanced Analytic Methods
  - Lesson 12: An Organized Approach to Analyzing a Dataset
- Module 7 – Evidence Synthesis
  - Lesson 13: Systematic Reviews and Meta-Analysis
  - Lesson 14, Part 1: Network Meta-Analysis
  - Lesson 14, Part 2: Example of a Published Network Meta-Analysis

- Module 8 – Potential Uses of AI and ML in HEOR
  - Lesson 15: Potential Uses of AI in HEOR
  - Lesson 16: Clinical Prediction Models
  - Lesson 17: AI-Driven Advancements in Evidence Synthesis: Transforming Systematic Reviews and Meta-Analyses

## **HEOR DS 520:**

### **Comparative Effectiveness Analysis with Routine Care Data**

Winter Quarter

Course 2 of 3 in the Data Science in HEOR Certificate Program

Instructor: Anirban Basu, PhD, [basua@uw.edu](mailto:basua@uw.edu)

#### **Course Description**

An introduction to causal inference methods. Causality is often considered the holy grail of empirical analysis. Researchers have approached causality from econometric, statistical, behavioral, and even philosophical points of view. This course will review statistical and econometric methods developed for individual-level data on behavior (choices or exposures) and outcomes to make causal inferences. In particular, the course will introduce a wide array of causal inference methods and will lay out the fundamental assumptions that researchers need to make and defend in order to make causal claims when applying a specific method to a particular type of data.

#### **Course Learning Outcomes**

By the end of this course, students will be able to:

- Understand introductory concepts in causality, how to choose among covariates, and regressions for outcomes
- Analyze health care expenditure data
- Explain regressions for utilizations, count data and survival outcomes
- Define methods for evaluating policy and estimating causal treatment effects and projections using real-world data
- Understand and analyze longitudinal or panel data

#### **Course Lessons**

- Module 1 – Introduction to Causal Inference
  - Lesson 1: Correlation is not Causation – Introduction to Causal Inference
  - Lesson 2: What Should I Control For? – Choosing Variables Wisely
  - Lesson 3: To Do or Not to Do? - Choosing Variables in Empirical Settings
  - Lesson 4: The Known Unknowns – Dealing with Unobserved Common Causes
  - Lesson 5: The Battle of Effects – Introduction to Alternate Target Parameters

- Module 2 – Outcomes-Based Regression Analysis
  - Lesson 6: Coin Flip – Regressions for Binary Outcomes
  - Lesson 7: Categories – Regressions for Multinomial Outcomes
  - Lesson 8: Fistful of Dollars – Analyzing Data on Healthcare Expenditures
  - Lesson 9: A Few Dollars More – Analyzing Data on Healthcare Expenditures
  - Lesson 10: Who is Counting? – Analyzing Count Data on Utilizations
  - Lesson 11: Survival of the Fittest – Introduction to Comparative Survival Analysis
  - Lesson 12: Sweeping Money Under the Rug – Analyzing Cost Data with Censoring
- Module 3 – Matching Estimators
  - Lesson 13a: Meet Your Match - Part I – Estimating Propensity Scores
  - Lesson 13b: Meet Your Match - Part II – Estimating Treatment Effects using Propensity Scores
  - Lesson 14a: Match Made in Heaven – Alternate Uses of Matching Methods - Part 1
  - Lesson 14b: Math Made in Heaven - Alternate Uses of Matching Methods - Part II
- Module 4 – Dealing with Unobserved Confounding
  - Lesson 15: IV Drip – Introduction to Instrumental Variable Analysis
  - Lesson 16: Mind the Gap – Introduction to Regression Discontinuity Designs
- Module 5 – Policy Evaluation with Panel or Longitudinal Data
  - Lesson 17: Long Data – Introduction to Panel/Longitudinal Data Modeling
  - Lesson 18: Pardon the Interruption – Introduction to Interrupted Time-Series Approaches
  - Lesson 19: DID You Do Two Subtractions? – Difference-in-Differences
  - Lesson 20: Synthetic Thinking – Introduction to Synthetic Control Approaches
- Module 6 – Inference
  - Lesson 21: Laces

## **HEOR DS 530:**

# **Principles of Machine Learning & Applications in Health Economics & Outcomes Research**

Spring Quarter

Course 3 of 3 in the Data Science in HEOR Certificate Program

Instructors:

Aastha Bansal, PhD, [abansal@uw.edu](mailto:abansal@uw.edu)

Noemi Kreif, PhD, [nkreif@uw.edu](mailto:nkreif@uw.edu)

## Course Description

An introduction to machine learning concepts and their application to the field of health economics and outcomes research. The course goal is to give students the skills to be able to understand the applied machine learning literature and communicate results to a general scientific audience. The course introduces a wide array of machine learning methods for prediction of outcomes and lays out the fundamental differences between different classes of methods, discusses the application of machine learning to causal inference, and ends with considerations for the real-world deployment of machine learning in healthcare, such as discrimination in health care decision-making with machine learning algorithms, model drift, generalizability, and regulation.

## Course Learning Outcomes

By the end of this course, students will be able to:

- Understand introductory concepts in machine learning
- Read and interpret articles that apply machine learning in HEOR
- Explain differences between supervised vs. unsupervised learning
- Understand introductory concepts in algorithmic fairness
- Describe considerations for the real-world application of machine learning in healthcare, such as model drift, generalizability, and regulation

## Course Lessons

- Module 1 – Introduction to Machine Learning
  - Lesson 1: Introduction to Machine Learning in HEOR
  - Lesson 2: Application of Machine Learning to Prediction vs. Inference
  - Lesson 3: Machine Learning Application Landscape in HEOR
- Module 2 – Supervised Learning for Regression Problems
  - Lesson 4: Evaluating Model Predictive Accuracy for Regression Problems
  - Lesson 5a: Introduction to Supervised Learning Methods for Regression Problems – Part I, Regression-Based Methods
  - Lesson 5b: Introduction to Supervised Learning Methods for Regression Problems – Part II, Tree-Based Methods
  - Lesson 5c: Introduction to Supervised Learning Methods for Regression – Part III, Super Learner and Deep Learning
- Module 3 – Supervised Learning for Classification Problems
  - Lesson 6: Evaluating Model Predictive Accuracy for Classification Problems
  - Lesson 7a: Introduction to Supervised Learning Methods for Classification Problems – Part I
  - Lesson 7b: Introduction to Supervised Learning Methods for Classification Problems – Part II

- Module 4 – Unsupervised Learning Methods
  - Lesson 8a: Introduction to Unsupervised Learning Methods - Part I, Principal Components Analysis
  - Lesson 8b: Introduction to Unsupervised Learning Methods - Part II, Clustering
- Module 5 – Machine Learning for Estimation of Treatment Effects
  - Lesson 9a: Machine Learning for Causal Inference: Average Treatment Effects – Part I
  - Lesson 9b: Machine Learning for Causal Inference: Average Treatment Effects – Part II
  - Lesson 9c: Machine Learning for Causal Inference: Average Treatment Effects – Part III
  - Lesson 10: Estimating Heterogeneous Treatment Effects
- Module 6 – Algorithmic Bias
  - Lesson 11: Introduction to Algorithmic Bias
  - Lesson 12a: Clinical Algorithmic Discrimination – Part I
  - Lesson 12b: Clinical Algorithmic Discrimination – Part II
- Module 7 – Considerations for the Real-World Application of Machine Learning in Healthcare and Other Contemporary Topics
  - Lesson 13: Informing Treatment Stratification Using Machine Learning
  - Lesson 14: Deployable Machine Learning for Healthcare
  - Lesson 15: Regulation of AI in Healthcare
  - Lesson 16: Generative AI in Healthcare: The Next Frontier