Clinical Sciences and Health Informatics (CSHI)

Course information contained within the Bulletin is accurate at the time of publication in June 2025 but is subject to change. For the most up-to-date course information, please refer to the Course Catalog.

CSHI 5201. Clinical Data and Analytics. 3 Credit Hours.
This course is designed to train students in the fundamentals of manipulating and conducting analyses of clinical data. Basic statistical principles and big data analytics will be discussed briefly, though the main focus of this course is understanding and working with "raw" clinical and administrative data sources to develop structured data sets suitable for analysis by more advanced techniques. Students will gain insight into the unique idiosyncrasies of the healthcare setting that impact the interpretation of analyses of clinical data. This class will explore in detail an anonymized patient data set including demographics, clinical encounters, diagnoses, medications, and laboratory results. The class will analyze clinical data sets with special focus on medication management, patient adherence, healthcare disparities, trends in health utilization, disease management, and quality measurement. Challenging issues in dealing with "dirty data" will be addressed. The course will also explore the impact of data modeling decisions and clinical ontology choices on the conclusions that can be drawn from clinical analytics. Through the course, students will learn the many assumptions that are incorporated into even seemingly straightforward queries, and the impact of these assumptions on the interpretation of results. Students also will be able to test the generalizability of their knowledge through analyses of additional publicly available clinical data sets.

Repeatability: This course may not be repeated for additional credits.

CSHI 5202. Applications of Data Science to Clinical Medicine and Medical Research. 2 Credit Hours.
Data science, at the intersection of statistics and computer science, provides cutting edge tools for analyzing large data sets to provide predictive and causal models. These tools have been used extensively in the fields of econometrics and advertising (e.g. Google and Facebook); the application of data science to medicine is in its infancy, however, with the increasing availability of large clinical data sets, the opportunity to mine these sets for information has never been greater. While software engineers and applied mathematicians have the skills to use and develop these tools, they lack the content expertise needed to apply this knowledge to medicine. The purpose of this course is to provide graduate student/researchers in biomedical sciences and clinical research staff an opportunity to develop sufficient skills in modern data science to carry out research using large databases. It is assumed that the student has no training beyond introductory courses in statistics and/or epidemiology; no significant programming experience is required.

Repeatability: This course may not be repeated for additional credits.