BIOSTATISTICS


Degree Types: MS

The Master of Science (MS) in Biostatistics program (https://www.feinberg.northwestern.edu/sites/cehs/our-programs/program-in-public-health/prospective-students) is a one-year program, providing graduate biostatistics training for students who intend to plan, direct and execute health research and/or analyze health data. The MS in Biostatistics program is distinguished by its concurrent emphasis on both statistical methodology and computer programming skills. It is a full-time program completed in four quarters or a half-time program completed in eight quarters. All students complete a thesis project in collaboration with biostatistics faculty and other Feinberg School of Medicine faculty researchers.

Three concentration options address a variety of student goals. The Concentration in Population Health Analytics is designed for college graduates or students with professional degrees (e.g. MD, DPT, allied health professionals) who intend to plan, direct and execute health research. The Concentration in Statistical Bioinformatics is designed for college graduates who are interested in working as statistical analysts/programmers on research teams and emphasizes cutting edge computation and analysis for genomics and other bioinformatics 'big data'. The Concentration in Statistical Methods and Practice is designed for college graduates who are interested in working as statistical analysts/programmers on research teams and encompasses a broader range of statistical theory and methods for data from health and medical research settings.

Upon completion of the program, students will be well-qualified for positions in academic research departments or in the pharmaceutical, insurance, or health care consulting industries.

Degrees Offered

• Biostatistics MS (https://catalogs.northwestern.edu/tgs/biostatistics/biostatistics-ms)

Biostatistics Courses

EPL_BIO 301-0 Introduction to Epidemiology (1 Unit)
This course introduces epidemiology and its uses for population health research. Concepts include measures of disease occurrence, common sources and types of data, important study designs, sources of error in epidemiologic studies and epidemiologic methods.

EPL_BIO 302-0 Introduction to Biostatistics (1 Unit)
This course introduces principles of biostatistics and applications of statistical methods in health and medical research. Concepts include descriptive statistics, basic probability, probability distributions, estimation, hypothesis testing, correlation and simple linear regression.

EPL_BIO 303-0 Basic Probability (1 Unit)
This course introduces probability as the theoretical framework underlying statistical methods. Concepts include random variables, discrete and continuous probability distributions, multivariate distributions, and random variable transformations. A working knowledge of differential / integral calculus and matrix algebra fundamentals is required.

EPL_BIO 305-0 Applied Statistical Programming in SAS (1 Unit)
This course provides a thorough working introduction to the statistical programming language SAS. Concepts focus on practical issues relating to data management, statistical data processing and SAS programming.

EPL_BIO 306-0 R Programming (1 Unit)
This course provides a thorough working introduction to the statistical programming language R. Concepts focus on practical issues including: installing and configuring the RStudio development environment; loading and managing data in R; accessing R packages; writing R functions; writing R scripts; debugging and profiling R scripts; organizing and commenting R code; and developing dynamic analysis reports using R Markdown. Topics in biostatistical data analysis will provide relevant working examples.

EPL_BIO 401-0 Intermediate Epidemiology (1 Unit)
The purpose of this course is twofold: 1) To elaborate on concepts first introduced in introductory epi and demonstrate to students how these concepts are frequently applied in biomedical literature; and 2) To provide students an overview of the physiology, pathophysiology, and epidemiology of prevalent diseases in the United States. Prerequisite: EPI_BIO 301-0, EPI_BIO 302-0 or PUB_HLTH 302-0.

EPL_BIO 402-0 Intermediate Biostatistics (1 Unit)
This course provides an intermediate-level treatment of linear and logistic regression models, including model estimation and inference, model building and diagnostics, and interpretation of results in the context of epidemiologic and clinical studies. The focus is on practical application of regression models for data analysis. The course uses R statistical software for data analysis.

EPL_BIO 403-0 Statistical Inference and Applied Techniques (1 Unit)
This course introduces statistical inference concepts and applied statistical techniques for data analysis in a mathematical framework. Concepts include point and interval estimation, maximum likelihood, large sample theory, hypothesis testing, bootstrap methods, analysis of variance, linear regression, analysis of categorical data and Bayesian methods.

EPL_BIO 427-0 Practical Issues in Population Studies (0.5 Unit)
This course covers practical aspects of conducting a population-based research study. Concepts include determining a study budget, setting a timeline, identifying study team members, setting a strategy for recruitment and retention, developing a data collection protocol and monitoring data collection to ensure quality control and quality assurance. Students will demonstrate these skills by engaging in a quarter-long group project to draft a Manual of Operations for a new "mock" population study.

EPL_BIO 429-0 Systematic Review and Meta-Analysis in the Medical Sciences (1 Unit)
This course covers statistical methods for meta-analysis. Concepts include fixed-effects and random-effects models, measures of heterogeneity, prediction intervals, meta regression, power assessment, subgroup analysis and assessment of publication bias. The course will emphasize basic theory and underlying statistical methods, computational approaches and interpretation of results from published studies.

EPL_BIO 445-0 Data Science for Clinical, Translational, and Population Researchers (1 Unit)
Due to advances in technology and data collection, the ability to analyze complex data sets is a necessary skill for all clinical, translational and population researchers. A variety of data analysis tools exist, some unique to specific domains. This course provides an introduction to the data, analysis tools, ethical considerations, and terminologies...
from across biomedical data science with an emphasis on clinical, translational and population methods and tools.

**EPI_BIO 446-0 Design, Conduct, and Analysis of Clinical Trials (1 Unit)**
This is an Introduction to clinical trials, including design, conduct, analysis and interpretation of trial results. Topics include trial design, methods for randomization, blinding and sample size determination, choice of controls, operational issues, data management and data quality issues, interim analysis methods, critical review of clinical trial proposals and published results and statistical techniques for analyzing clinical trial data.

**EPI_BIO 499-0 Independent Study (0.5-1 Unit)**
This is Independent Study course is designed to give students a hands-on practical training on statistical analysis. This course will offer 4 statistical consultation projects -- real case study. Students will be given the data and a brief introduction of the data. Students will need to check the data, perform preliminary data analysis, orally present their statistical approach, modeling, findings, and interpretation of their result, and write a report for each project. Students will be evaluated by the quality of their oral presentation and project report.

**EPI_BIO 501-0 Advanced Epidemiology (1 Unit)**
This course builds on material learned in previous Biostatistics and Epidemiology courses. Concepts are applied to the design, implementation, analysis and interpretation of observational epidemiologic studies (cross-sectional, case-control and cohort). Students enrolled in an MPH degree program must have the consent of the instructor.
Prerequisites: EPI_BIO 301-0, EPI_BIO 302-0, EPI_BIO 401-0, EPI_BIO 402-0.

**EPI_BIO 502-0 Advanced Biostatistics (1 Unit)**
This course covers modern approaches to the analysis of correlated response data arising from longitudinal studies commonly encountered in medical research and clinical trials. Concepts include marginal and mixed-effects regression models for continuous and discrete outcomes measured repeatedly over time, model building techniques, robust inference procedures and problems associated with missing data. All modeling and numerical analyses will be done using SAS. Prerequisites include the equivalent of EPI_BIO 302-0, EPI_BIO 402-0.

**EPI_BIO 521-0 Applied Survival Data Analysis (1 Unit)**
This courses provides an introduction to the fundamental concepts and methods developed for analysis of survival data for which incompleteness, including censoring, is a primary feature. Classic non-parametric estimation approaches will be discussed, as will semi-parametric and parametric hazard regression modeling techniques that allow incorporation of covariates. Analysis examples using both R and SAS will be discussed.

**EPI_BIO 527-0 Statistical Methods for Missing Data (0.5 Unit)**
This course provides students with a basic knowledge of the potential implications of missing data on their data analyses as well as potential solutions. A major focus of the course is multiple imputation including discussions of the general framework, different models and algorithms, and the basic theory. Statistical programming is performed in R.

**EPI_BIO 529-0 Statistical Genetics (0.5 Unit)**
This course equips students with key principles and practical skills to analyze genetic data. Topics range from linkage analysis using pedigree data to machine learning techniques using next-generation sequencing data. Statistical programming is performed in R.

**EPI_BIO 560-0 Statistical Consulting (0.5 Unit)**
This course prepares students for collaboration and communication with scientists of various disciplines, emphasizing analytical tools, verbal and written communication skills and presentation skills. Concepts include sample size and power calculation, handling of various data structures, data presentation, selecting appropriate statistical methods, time and project management, reproducible research, report writing and grant writing.

**EPI_BIO 561-0 Thesis (0.5 Unit)**
All MS in Biostatistics students are required to submit a master's thesis focusing on a collaborative data analysis or a statistical methodology advancement. Thesis projects are proposed in the fall, and content is developed primarily in late fall / winter. Written, poster and oral presentations are due in the spring.

**EPI_BIO 565-0 Clinical Database Management (0.5 Unit)**
This course serves as an introduction to data management in the clinical research setting. The Research Electronic Data Capture (REDCap) platform is used to understand basic database design, data management and quality monitoring concepts for studies ranging from simple cross-sectional designs to complex multi-center clinical trials.