BIOSTATISTICS

Degree Types: MS

The Master of Science (MS) in Biostatistics program (https://www.feinberg.northwestern.edu/sites/cehs/our-programs/master-of-science-in-biostatistics/) 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 Population Health Analytics concentration 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 Statistical Bioinformatics concentration 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 Statistical Methods and Practice concentration 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.

Additional resources:

- Department website (https://www.feinberg.northwestern.edu/sites/cehs/our-programs/master-of-science-in-biostatistics/)
- Program handbook(s)

Degrees Offered

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

Learning objective(s)/Students should be able to…

- Apply classic methods for continuous and categorical data analysis, including regression and other appropriate statistical approaches
- Use computer-based statistical analysis package(s) to manage data
- Develop visualized data using computer-based statistical analysis package(s)
- Analyze data employing computer-based statistical analysis package(s)
- Implement sample size and power calculations for a range of experimental designs
- Interpret results of a health research study, including the relation to findings from other studies, potential biological or social mechanisms, study limitations, and public health implications
- Communicate findings in a scientifically sound manner in written and verbal format
- Calculate epidemiological measures of association between risk factors and disease
- Apply methods and strategies to evaluate and reduce bias in health research
- Use criteria to distinguish between association and causality
- Apply ethical and regulatory standards to human subjects' research
- Explain public health history, philosophy and values.
- Identify the core functions of public health and the 10 Essential Services
- Explain the role of quantitative and qualitative methods and sciences in describing and assessing a population's health
- List major causes and trends of morbidity and mortality in the US or other community relevant to the school or program
- Discuss the science of primary, secondary & tertiary prevention in population health, including health promotion, screening, etc.
- Explain the critical importance of evidence in advancing public health knowledge
- Explain effects of environmental factors on a population's health
- Explain biological and genetic factors that affect a population's health
- Explain behavioral and psychological factors that affect a population's health
- Explain the social, political and economic determinants of health and how they contribute to population health and health inequities
- Explain how globalization affects global burdens of disease
- Explain an ecological perspective on the connections among human health, animal health and ecosystem health (e.g., One Health).

Concentration in Population Health Analytics

- Design an epidemiologic study to address a question of interest
- Describe practical considerations for the conduct of health research studies
- Access publicly available data resources for population health research
- Critically review the scientific literature, synthesize findings across studies, and make appropriate recommendations based on current knowledge
- Develop a clear description of the rationale, methods, results, and overall interpretation of an epidemiologic investigation

Concentration in Statistical Bioinformatics

- Develop computer files of high-dimensional data for analysis using high performance computing data management techniques
- Determine and execute appropriate statistical analyses, in particular techniques relevant to bioinformatics, to address a study question
- Access publicly available databases for bioinformatics research
- Develop statistical and bioinformatics analysis results in written, graphical and verbal format in response to an analysis request
- Identify theoretical underpinnings of advanced statistical models

Concentration in Statistical Methods and Practice

- Develop computer files of raw data for analysis using data management and statistical analysis software

- Execute appropriate statistical analyses to address a study question
- Apply classic methods for the analysis of time-to-event and clinical trial data
- Develop statistical analysis results in written and verbal format in response to an analysis request
- Identify theoretical underpinnings of advanced statistical models

Biostatistics Courses

BIOSTAT 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.

BIOSTAT 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.

BIOSTAT 303-0 Introduction to Probability Theory (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.

BIOSTAT 305-0 Data Management and Analysis 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. Prerequisites: BIOSTAT 302-0.

BIOSTAT 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.

BIOSTAT 390-0 Emerging Topics in Biostatistics (1 Unit)
This course is student interest-driven; the student(s) will each select topic(s) related to clinical trials that they would like to understand in greater depth. The course will function much like a journal club, and each student will present on his/her specified topics. Those not presenting will be responsible for review and critique of the discussion.

BIOSTAT 401-0 Intermediate Epidemiology (1 Unit)
The purpose of this course is to elaborate on concepts introduced in introductory epi, demonstrate how these concepts are frequently applied in biomedical literature, and to provide an overview of the physiology, pathophysiology, and epidemiology of prevalent diseases in the US. Prerequisite: BIOSTAT 301-0.

BIOSTAT 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. Prerequisites BIOSTAT 302-0 and BIOSTAT 306-0.

BIOSTAT 403-0 Statistical Inference (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. A working knowledge of college calculus I and II is essential. Prerequisite: BIOSTAT 303-0.

BIOSTAT 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.

BIOSTAT 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. Prerequisite BIOSTAT 302-0.

BIOSTAT 446-0 Design, Conduct and Analysis of Clinical Trials (1 Unit)
This course introduces commonly used designs for clinical trials, methods for randomization, blinding and sample size determination, choice of controls, collaborative/multicenter trial requirements and operational issues, data management and data quality issues, interim analysis methods, critical review of clinical trial results and statistical techniques for analyzing data. Prerequisite BIOSTAT 305-0 and BIOSTAT 402-0.

BIOSTAT 499-0 Independent Study (0.5-1 Unit)
The Independent Study provides practical training using statistical consultation projects. After a brief introduction of the data, students will check the data, perform preliminary data analysis, orally present their statistical approach, modeling, findings, and interpretation of the result, and write a report. Students will be evaluated by the quality of their oral presentation and project report.

BIOSTAT 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 a degree program outside of the MS in Biostatistics must have the consent of the instructor.
Prerequisites: BIOSTAT 305-0, BIOSTAT 401-0, BIOSTAT 402-0.

BIOSTAT 502-0 Advanced Biostatistics (1 Unit)
This course covers analysis of correlated response data arising from longitudinal studies. 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: BIOSTAT 302-0, BIOSTAT 305-0, BIOSTAT 402-0, or instructor approval.

BIOSTAT 521-0 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.
Prerequisite: BIOSTAT 403-0.

BIOSTAT 522-0 Network Data Analysis (0.5 Unit)
This course introduces analytic methods and computational resources for network data analysis. It provides an overview dealing computationally with graph and network data structures, describing networks through node and edges statistics and overall topology, and identifying cliques and clusters in graphs. Gaussian graphical models, subnetwork identification and differential network analysis will be introduced.
Prerequisites: BIOSTAT 302-0, BIOSTAT 306-0, BIOSTAT 402-0.

BIOSTAT 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.
Prerequisites: BIOSTAT 306-0, BIOSTAT 402-0, BIOSTAT 403-0.

BIOSTAT 529-0 Statistical Genetics and Genomic Data Analysis (1 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.
Prerequisites: BIOSTAT 306-0, BIOSTAT 403-0, or instructor approval.

BIOSTAT 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.
Prerequisites: BIOSTAT 302-0 and BIOSTAT 402-0.

BIOSTAT 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.

BIOSTAT 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. Students enrolled in this course must complete CITI training through the Northwestern University IRB Office before the first class session.