Biostatistics

Degrees Offered
Doctor of Philosophy
Master of Science

Doctor of Philosophy Program

Program Admissions Requirements

*In addition to the general **Graduate School admission requirements**, this program has additional specific requirements.*

Any graduate of an accredited college or university with an undergraduate degree in mathematics or closely related fields with strong preparation in mathematics is eligible for admission. Applicants are expected to have completed and performed well in courses in advanced calculus, linear/matrix algebra and scientific programming. Those who haven't done so may be considered for admission to the program upon approval of the biostatistics admission committee, and if admitted, these requirements must be completed during the first year of study. In addition, the applicant must have strong interest in pursuing statistical research in biomedical sciences.

Fields of Research

- Markov chain Monte Carlo, statistical methods for epidemiology and biology.
- Bayesian statistics, linear and generalized linear models, nonparametric inference and Monte Carlo methods.
- Multiple comparisons, dose-response studies, clinical trials, multiple endpoints, fMRI/image data analysis.
- Categorical and correlated data analysis, statistical models in epidemiology.
- Estimation with additional information, missing data, censored data, survey data.
- Statistical genetics, linkage and association mapping, haplotype analysis
- Bioinformatics.
- Survival and competing risks data analysis, inference from stochastic processes and non-linear models.
- Personalized medicine, causal inference
- Machine learning, high dimensional data analysis, variable selection

Overall Course Requirements
A requirement of this program is to fulfill two credits in Bioethics by completing Course (10222) Ethics and Integrity in Science and Course (10444) Research Ethics Discussion Series. For course descriptions of 10222 and 10444 see listing within the **Bioethics Program**.

** Indicates University of Wisconsin Milwaukee (UWM) course number. All UWM courses require an off-campus course registration procedure.*

PhD Degree Required Courses

10222 Ethics and Integrity in Science. *1 credit.*
10444 Research Ethics Discussion Series. *1 credit.*
04214 Design and Analysis of Clinical Trials. *3 credits.*
04220 Research Seminar. *1-9 credit(s).*
04221 Biomedical Applications and Consulting *3 credits.*
04222 Statistical Consulting. *3 credits.*
04224 Biostatistical Computing. *3 credits.*
04231 Statistical Models and Methods I. *3 credits.*
04232 Statistical Models and Methods II. *3 credits.*
04233 Introduction to Statistical and Machine Learning. *3 credits.*
04275 Applied Survival Analysis. *3 credits.*
04285 Introduction to Bayesian Analysis. *3 credits.*
04295 Reading and Research. *1-9 credit(s).*
04313 Advanced Statistical Computing. *3 credits.*
04363 Advanced Statistics I. *3 credits.*
04365 Linear Models I. *3 credits.*
04384 Statistical Genetics. *3 credits.*
04385 Advanced Bayesian Analysis. *3 credits.*
04386 Theory of Survival Analysis. *3 credits.*
04399 Doctoral Dissertation. *1-9 credits.*

04231/MTHSTAT 761* Mathematical Statistics I. *3 credits.*
04232/MTHSTAT 762* Mathematical Statistics II. *3 credits.*
PH721* Introduction to Translational Bio-informatics. *3 credits.*
*courses taken at UW-Milwaukee

A minimum of six credit hours of electives in a non-statistical field such as biological/medical science, mathematics, and computer science is a requirement for a candidate seeking the PhD degree in Biostatistics. Examples of courses meeting this requirement are:

BIOETH 201 Medical Ethics.	*2 credits.*
BIOETH 222 Ethics and Integrity in Science.	*2 credits.*
BIOETH 232 Ethics, Policy and Genetic Technology.	*2 credits.*
BIOPHYSICS 215 Medical Physics.	*1 credit.*
CELLBIO 150 Introductory Cell Biology.	*1 credit.*
CELLBIO 152 Human Development.	*1 credit.*
CELLBIO 207 Introduction to Neuroscience.	*2 credits.*
EPI 201 Clinical Epidemiology.	*3 credits.*
EPI 256 Research Methods in Epidemiology.	*3 credits.*
EPI 272 Epidemiology of Cardiovascular Disease.	*1 credit.*
EPI 274 Cancer Epidemiology.	*1 credit.*
PHARM 202 Survey of Pharmacology.	*3 credits.*
PHY 285 Mathematical Biology.	*3 credits.*

Examination Process for Ph.D. degree

A. **Preliminary Examinations:** Upon completion of the necessary courses, the student is given two written preliminary examinations. One addresses Theory of Statistics, covering the subject matter from Statistical Inference I & II. The other, on Applied Statistics, covers Statistical Models and Methods I, II and III, Design and Analysis of Clinical Trials, Applied Survival Analysis, Introduction to Bayesian Analysis, and Biostatistical Consulting. Both examinations are organized and administered by the Division's Graduate Studies Committee. Evaluation is done by the entire faculty. The criteria for evaluation are the student's understanding and competency in basic principles and foundations of biostatistics, and his/her potential for conducting independent research in statistical methods and applications. To continue in the Ph.D. program, both examinations must be successfully completed by the end of August in the student's second year. The examinations are offered every August. If a student does not pass an exam, he/she is given a second opportunity to take it in January.

B. **Choosing an Advisor and forming a Dissertation Committee:** By the beginning of the Fall semester in the student's third year in the program, the student chooses a member of the Division's faculty as his/her advisor with mutual agreement. It is expected that this choice will grow out of the student's coursework, two research and readings courses, seminar participation and general immersion in the Division's academic activities. The two readings courses are taken in the first summer and the second spring or summer

semesters. The student's choice of advisor must be approved by the Director of Graduate Studies and the Division Director. Advisors are typically chosen from the senior faculty. A junior faculty member who has not previously served as dissertation advisor may do so provided a more experienced faculty member is chosen as co-advisor.

In close consultation with the advisor, the student forms the Dissertation Committee in full accordance with the requirements of the Graduate School. The committee consists of five graduate faculty members including the advisor. Four of the five must be from the Division of Biostatistics (including Joint and Adjunct faculty) and one must be from outside the Division of Biostatistics. The committee must be approved by the Director of Graduate Studies and the Division Director. The process of committee formation, including submission of the appropriate form to the Graduate School, must be completed by the end of September in the student's third year. From this date forward the student's progress is monitored by the advisor and the Dissertation Committee.

C. Qualifying Examination: Upon successful completion of the preliminary exam and at a time determined by the Dissertation Committee, the student is given a qualifying examination. This examination is individualized for each student, and it is organized, administered and evaluated by his/her Dissertation Committee. The evaluations are based on student's in-depth understanding and competency in advanced topics in biostatistics, and his/her ability and maturity to apply the knowledge earned from the course-work in conducting meaningful research. The exam consists of two parts. One part is an oral examination testing the student's general statistical knowledge at the advanced level. The other part consists of writing a dissertation proposal and presenting it to the Division. This proposal must be approved by his/her Dissertation Committee. A student not passing either part of the exam may be given another chance to retake that part within three months of the first attempt. Students passing this exam will be admitted to Ph.D. candidacy.

D. Paper submission: The student is required to submit at least one methodology paper to peer reviewed journals. The paper must address statistical methodology and be from the

thesis. The student must provide a proof of paper submission for the thesis committee before the final examination.

E. Final Examination: The PhD candidate must submit a dissertation representing an original research contribution. It must show high attainment and clear ability to carry out independent biostatistics research of publishable quality. The final oral examination, including a public defense of the dissertation, is administered by his/her Dissertation Committee after the student has completed all other formal requirements for the PhD degree. The student is expected to demonstrate a good understanding of the general field in which the dissertation is written. The student's Dissertation Committee will evaluate the performance of the student in the dissertation defense.

Master of Science Program

Required Courses

All of the following courses:

BIOST 04214 Design and Analysis of Clinical Trials. *3 credits.*
BIOST 04224 Biostatistical Computing. *3 credits.*
BIOST 04231 Statistical Models and Methods I. *3 credits.*
BIOST 04232 Statistical Models and Methods II. *3 credits.*
BIOST 04275 Applied Survival Analysis. *3 credits.*
BIOST 04285 Introduction to Bayesian Analysis. *3 credits.*
BIOST 04221 Biomedical Applications and Consulting *3 credits.*
BIOST 04222 Statistical Consulting. *3 credits.*
BIOETH 1022a Ethics and Integrity in Science. *1 credit.*
04231/MTHSTAT 761* Mathematical Statistics I. *3 credits.*
04232/MTHSTAT 762* Mathematical Statistics II. *3 credits.*

Any three of the following courses:

BIOST 04264* Time Series Analysis. *3 credits.*
BIOST 04280* Applied Probability. *3 credits.*
BIOST 04232 Introduction to Statistical and Machine Learning. *3 credits.*
BIOST 04313 Advanced Statistical Computing. *3 credits.*
BIOST 04363 Advanced Statistics I. *3 credits.*
BIOST 04365 Linear Models I. *3 credits.*
BIOST 04384 Statistical Genetics. *3 credits.*
BIOST 04385 Advanced Bayesian Analysis. *3 credits.*
BIOST 04386 Theory of Survival Analysis. *3 credits.*
PH721* Introduction to Translational Bio-informatics.

3 credits.
*Courses taught at UWM Mathematics Department

Electives
Any two of the following:
MTHSTAT 564A * Time Series Analysis. *3 credits.*
MTHSTAT 571 * Introduction to Probability Models. *3 credits.*
PH721 * Introduction to Translational Bioinformatics. *3 credits.*
04313 Advance Statistical Computing. *3 credits.*
04363 Advance Statistics I. *3 credits.*
04365 Linear Models I. *3 credits.*
04384 Statistical Genetics. *3 credits.*
04385 Advanced Bayesian Analysis. *3 credits.*
04386 Theory of Survival Analysis. *3 credits.*

Writing Requirement
Students are required to prepare written reports on two consulting/collaborative research projects. These reports should include a description of the biological problem, a discussion of the statistical methods used in the analysis and a presentation of results. The reports must be written for presentation to the clinical investigator and not be focused solely on statistical techniques. A guide to writing consulting reports can be found in The Statistical Consultant in Action by DJ Hand & BS Everitt, Cambridge University Press, 1987. Reports can be based either on projects from the student's consulting classes or from the student's work assignment. The papers should be 5-10 pages in length as a guide. The documents must be approved by a faculty member (typically the instructor of the consulting class or a member of the student's examination committee).

Courses
04200 Biostatistics I. *3 credits.*
This is an introductory course in biostatistical methods for non-biostatistics majors. Topics include elementary probability, sampling, point and interval estimation and hypothesis testing.
04201 Biostatistics II. *3 credits.*
A continuation of Biostatistics I. Topics include statistical methods for categorical data, regression and correlation, and analysis of variance.
04202 Principles of Biostatistics. *1 credit.*
This course provides an introduction to statistical concepts used in medical research at a non-mathematical level. Topics include introduction to study designs, descriptive statistics, probability, estimation, test of hypothesis, regression and correlation.
04214 Design and Analysis of Clinical Trials. *3 credits.*

Prerequisites: Statistical Models and Methods I or concurrent registration
This course covers issues in clinical trials including the clinical trial protocol, sources of bias in clinical trials, blinding, randomization, sample size calculation; phase I, phase II, phase III and hybrid trials; interim analysis, stochastic curtailment, Bayesian designs, and administrative issues in study design.

04220 Research Seminar. *1 credit.*
Prerequisites: Concurrent registration
Students present plans for an analysis of research projects and research data. Projects and examples from classical and current literature are discussed by students and faculty.

04221 Biomedical Applications and Consulting. *3 credits*.
Prerequisites: Statistical Models and Methods I
Theory of consulting, communication and statistical techniques most often used in consulting and biomedical applications, practical experience in the real consulting setting and writing statistical reports.

04222 Statistical Consulting. *1-3 credit(s).*
Prerequisites: Statistical Models and Methods I & II
This course is designed for students to gain experience in statistical consulting by working with the biostatistics faculty members on various consulting projects.

04224 Biostatistical Computing. *3 credits.*
Prerequisites: Statistical Models and Methods I or concurrent registration
This course will cover the details of manipulating and transforming data required for statistical analysis. Topics include reshaping the data from a per-case to a per-event within a case and vice-versa. It will also cover the techniques necessary to write functions and macros in both SAS and R for developing new/modified data analysis methods. How to use R packages and C/C++ codes in R will also be covered. The LaTeX document production system is also introduced.

04231 Statistical Models and Methods I. *3 credits.*
Prerequisite: Three semesters of calculus and one semester of linear algebra
This course will cover statistical techniques for basic statistics. Topics include one-sample/two-sample tests, analyses for count data and contingency tables, basic nonparametric methods including sign, rank-sum and signed-rank tests, simple linear regression model and inference, checking model assumptions, model diagnostics, correlation analysis, one-way analysis of variance, Kruskal-Wallis one-way ANOVA, simple logistic regression, and weighted linear regression. SAS/R will be used throughout the course.

04232 Statistical Models and Methods II. *3 credits.*
Prerequisite: Statistical Models and Methods I
This course will cover various regression models for independent and correlated data. Topics include multiple linear regression, model diagnostics, variable selection, influence/leverage, outliers, collinearity, transformation, GLM including logistic and Poisson regression, overdispersion, GEE, mixed models, and GLMM. SAS/R will be used throughout the course.

04233 Introduction to Statistical and Machine Learning. *3 credits.*
Prerequisite: Statistical Models and Methods II
This course will provide an introduction to statistical learning. Core topics include variable selection, penalized linear regression such as lasso, dimension reduction including principal component analysis, flexible regression techniques including kernel smoothing/smoothing splines/generalized additive models/regression trees, support vector machine, clustering, and random forests. Other topics that can be covered include but are not limited to ridge regression, group lasso, fused lasso, adaptive lasso, SCAD, Bayesian lasso, Bayesian group lasso, Bayesian CART, BART, neural network, feature screening, graphical models, and quantile regression.

MTHSTAT 761 * Mathematical Statistics I. *3 credits.* (UWM registration)
Fundamentals of probability, independence, distribution and density functions, random variables, moments and moment-generating functions, discrete and continuous distributions, exponential families, location and scale families, marginal and conditional distributions, transformation and change of variables, multivariate distributions, random samples, convergence concepts, sampling from normal distributions, order statistics.

MTHSTAT 762 * Mathematical Statistics II. *3 credits.* (UWM registration)
Point estimation, interval estimation, hypothesis testing, minimal sufficiency and completeness, ancillary statistics, likelihood and invariance principle, asymptotic properties of estimators and likelihood ratio tests, LMP tests, union-intersection tests, pivotal quantities, coverage probability, large-sample estimation and testing.

PH721 * Introduction to Translational Bioinformatics. *3 credits.* (UWM registration)
Bioinformatics has become one of the major disciplines in modern biomedical research. Knowledge and analytic skills to retrieve the most relevant information imbedded in the large omic data are key to the discovery for translational research. This course will review high-throughput technologies that produce various omic data, along with

the methodologies and tools to analyze and interpret these different layers of information. Topics will cover a variety of data mining techniques and the use of several widely-used bioinformatics software and programming tools with emphasis on guiding students through the process of translating genomics data into biological knowledge, towards the discovery of novel therapeutic targets, biomarkers and the dissection of gene networks and pathways.

04275 Applied Survival Analysis. *3 credits.*

Prerequisites: Statistical Models and Methods I

The following topics will be covered in this course: Basic parameters in survival studies; Censoring and truncation, Competing risks; Univariate estimation including the Kaplan-Meier and Nelson-Aalen estimator; tests comparing two or more populations, the log rank test; Semi-parametric regression, the Cox model; Aalen's Additive hazards regression model; regression diagnostics.

04285 Introduction to Bayesian Analysis. *3 credits.*

Prerequisites: Statistical Models and Methods I

This course introduces basic concepts and computational tools for Bayesian statistical methods. Topics covered include one and two sample inference, regression models and comparison of several populations with normal, dichotomous and count data.

04295 Reading and Research. *1-9 credit(s).*

Prerequisites: Concurrent registration

Readings in recent literature and supervised research project.

04313 Advanced Statistical Computing. *3 credits.*

Prerequisites: Statistical Models and Methods II, Statistical Inference II, Biostatical Computing

This course will focus on numerical computing of statistics and algorithm programming. Topics include: numerical random number generation, likelihood maximization, numerical integration using quadrature and Monte-Carlo methods, the EM algorithm, Monte Carlo simulation, resampling (Bootstrap, permutation, Jackknife), optimization for penalized regression, parallel computing, and creating R packages.

04363 Advanced Statistics I. *3 credits.*

Prerequisites: Statistical Models and Methods II, Statistical Inference II

This course covers both the theoretical framework and practical aspects of statistical models. The course will cover likelihood inference, properties of likelihood, exponential families and GLM, large sample properties of likelihood-based inference, likelihood based regression models, GEE, conditional and marginal likelihood, asymptotics of penalized regression.

04365 Linear Models I. *3 credits.*
Prerequisites: Statistical Inference II
This course will cover review of matrix algebra and vector spaces, multivariate normal distribution, quadratic forms, least squares estimation, ANOVA, testing contrasts, multiple comparison, lack-of-fit test, multiple regression models, and mixed models. Emphasis is on theory.
04384 Statistical Genetics. *3 credits.*
Prerequisites: Linear Models I, Statistical Inference II
This course will cover the fundamental concepts in population genetics and statistical models and methods on genetic linkage and association mapping studies. Topics include Mendelian inheritance, Hardy-Weinberg equilibrium, linkage disequilibrium, allele identity by descent (IBD), inbreeding and coancestry coefficients, genetic models, heritability, genetic variance components, linkage analysis, haplotype analysis, case-control association analysis, association analysis using family data, adjust for population admixture, analysis of rare variants, microarray data analysis, eQTL analysis, copy number variants (CNV), RNA-seq data, proteomic and methylation data analysis.
04385 Advanced Bayesian Analysis. *3 credits.*
Prerequisites: Introduction to Bayesian Analysis
A combination of Bayesian principles, tools and methods; emphasis is on models, computations and analysis. Likelihood function, prior, posterior and predictive distributions, Bayes factors, HPD regions, conjugate and non-informative priors in the exponential family, Markov chain Monte Carlo methods for the generalized linear model, hierarchical models, restricted parameter spaces and censored data, examples of Bayesian analyses of complex biomedical models.
04386 Theory of Survival Analysis. *3 credits.*
Prerequisites: Applied Survival Analysis, Statistical Inference II
Analysis of survival data using counting process techniques. Topics include the mathematical theory of counting process, martingales, asymptotic properties for estimation of the survival and cumulative hazard functions, proportional hazards and additive hazards regression models, multivariate survival data, and empirical process.
04391 Special Topics in Statistics. *1-3 credit(s).*
Prerequisites: Concurrent registration
This course is designed to cover special topics in biostatistics that are not covered in regular courses. The topics will depend on the research interests of the instructor and the students.
04399 Doctoral Dissertation. *1-9 credit(s).*
Prerequisites: Concurrent registration