MPH/MSPH Degree Programs

Student Handbook

EMORY
ROLLINS SCHOOL OF PUBLIC HEALTH
Department of Biostatistics and Bioinformatics

http://www.sph.emory.edu/bios

THE FIELD OF BIOSTATISTICS

Biostatistics is the science that applies statistical theory and methods to the solution of problems in the biological and health sciences. A few examples of research questions which biostatistics can help answer are: What mathematical models can describe transmission and infection probabilities of infectious diseases such as AIDS and influenza? What are the risk factors associated with breast cancer? What preventive steps can people take to reduce the risk of heart disease? How many IV drug users have contracted AIDS in Georgia?

The main areas of effort for biostatisticians include collaborative research and consulting, methodological research, and education. In collaborative research, biostatisticians work on research studies with experts in the biological and health sciences. The biostatisticians' responsibilities include analysis of data and interpretation of results. Equally important, however, is the responsibility to assist in the planning and conduct of the study to ensure consistency with good statistical practice. Methodological research, such as developing mathematical models to describe biological phenomena, is conducted to enhance the existing bodies of knowledge in theoretical and applied biostatistics. Biostatisticians educate others about biostatistics through the teaching of graduate and continuing education courses, seminars, collaborative research and consulting activities.

Students entering graduate programs in biostatistics come from a variety of undergraduate fields. Many have undergraduate degrees in mathematics, applied mathematics or statistics. Others may have majored in the biological or social sciences. While specific requirements vary depending on the particular degree sought by a student, all students are expected to have a strong undergraduate background in mathematics or statistics, and a strong desire to study the theory and application of statistical methods in the biological and health sciences.

Upon enrolling in a biostatistics program, students take courses in statistical methods and theory. The methods courses focus on ways to select and apply statistical techniques that are appropriate for different types of problems. The theory courses provide rigorous instruction in the formal mathematical structure underlying the statistical techniques. Heavy use is made of computers in most biostatistics courses. Required and elective courses from other public health or biomedical fields are also included in the program of study.

Employment prospects for Masters level biostatisticians have been excellent in recent years. Positions as researchers and data analysts are commonly available in industry (e.g., pharmaceutical, consulting), academia (e.g., schools of public health and schools of medicine) and government agencies (e.g., the Centers for Disease Control and Prevention, local or state health departments). The monthly news magazine of the American Statistical Association (ASA), *Amstat News*, contains nationwide listings of career opportunities for biostatisticians. For more information about careers in biostatistics, visit the ASA website at www.amstat.org/careers.

THE DEPARTMENT OF BIOSTATISTICS AT EMORY UNIVERSITY

The Department of Biostatistics at Emory University was established in the early 1960's as the Department of Statistics and Biometry in the School of Medicine. In Fall 1990, it became one of six departments in the Rollins School of Public Health. The Department has 39 faculty, 15 adjunct and visiting faculty and one Emerita faculty.

Mission

The Department's mission is to develop, apply and provide training in quantitative methods that can improve human health by teasing out the information contained within modern data. This includes pursuing excellence in the four core responsibility areas:

> **Education:** Educate others about biostatistics through mentoring of and teaching graduate students, inter-disciplinary courses, continuing education courses, and seminars.
>
> **Methodological Research:** Conduct methodological research to enhance the existing bodies of knowledge in theoretical and applied statistics.
>
> **Collaborative Research:** Conduct collaborative research studies that use biostatistical methods with experts in the biological and health sciences in which the statistician makes substantial contributions, from assistance in the planning and conduct of the study to analysis of data and interpretation of results.
>
> **Service:** Provide statistical support for research projects outside the department that are limited in time, nature, and scope. Participate on local and national committees or other "citizenship" responsibilities.

Teaching

The Department of Biostatistics and Bioinformatics offers the MSPH and MPH degrees in Biostatistics and Public Health Informatics through the Rollins School of Public Health. The Department has approximately 100 students in our masters programs.

Research

The research activities of the faculty are diverse and include studies of national and international scope. Department faculty conduct research relating to survival analysis, biomedical imaging statistics, spatial statistics, geographic information systems, disease ecology, clinical trials, the analysis of missing and mismeasured data, next-generation sequencing, genomics, proteomics, longitudinal analysis, high-dimensional variable selection, and estimating equation theory. Department faculty also regularly collaborate in statistical issues relating to cardiology, ophthalmology, neurology, environmental epidemiology, reproductive epidemiology, aging, and quality of life.

Consulting/Collaboration

Faculty of the department collaborate with researchers at the Centers for Disease Control and Prevention, the Carter Center of Emory University, the Emory University School of Medicine, and other health-related organizations. The Department coordinates the activities of the Biostatistics Collaboration Core, which serves as a resource for advice on the design, conduct, and analysis of studies in the health sciences. Students may gain hands-on experience in practical biostatistical problems through working with faculty on real-life consulting problems.

Opportunities for Practical Training

Opportunities for practical experience are often available in the form of collaborative research with a Biostatistics and Bioinformatics faculty member or summer internships with agencies affiliated with the Department. Some of these opportunities provide limited stipends. In the past, students have held internships at the Centers for Disease Control and Prevention, and at pharmaceutical and consulting firms in the Atlanta area. Teaching opportunities have also been available.

REAL Program

Rollins Earn And Learn (REAL) is a **work-study program** funded by Rollins that offers eligible, full-time MPH and MSPH students valuable opportunities to *earn while they learn* through applied public health experiences in real-world settings. Opportunities include federal, state, county, and other government agencies, as well as Emory-affiliated programs, for-profit, and nonprofit organizations throughout Atlanta. This program helps approximately 500 Rollins students find meaningful public health work opportunities each year. These integral experiences often fulfill practicum requirements, lead to thesis opportunities, and provide an enriching experience for both employers and students. In other words, it's a win-win.

Computing

Faculty and students have access to a variety of central servers, workstations, and microcomputers running UNIX, Linux, Mac OS, Windows and other operating systems. Software includes EPIINFO, Fortran, IMSL, LaTeX, Mathematica, Minitab, R, SAS, SPSS, Splus, StatExact, , Sudaan, C, and C++, WinBUGS[1]. Communication capabilities include access to the Internet. The Department has its own computer lab for students to use.

[1] *The following products are trademarked by their respective companies: , UNIX, Linux, MacOS, Windows EPIINFO, Fortran, IMSL, LaTex, Mathematica, Minitab, R,SAS, SPSS, Splus, StatExact, , C, C++, and Sudaa, WinBUGS..*

AFFILIATED SCHOOLS, CENTERS & INSTITUTES

The Emory School of Medicine

The Emory School of Medicine is involved in an extensive program of teaching, research, and service. The School strives to offer the best possible learning opportunities in clinical medicine and research programs. Biostatistics and Bioinformatics faculty have extensive collaborative ties with researchers in the School of Medicine, including faculty at the General Clinical Research Center (GCRC), Winship Cancer Institute, and Departments of Genetics, Ophthalmology, Radiation Oncology, Pathology, Cardiology, Neurology, Rehabilitation Medicine, the Vaccine Center, and the Yerkes Primate Research Center.

The Centers for Disease Control and Prevention (CDC)

The CDC is a branch of the US Department of Health and Human Services that is internationally renowned for its work in public health. Biostatistics and Bioinformatics faculty have strong collaborative ties with researchers at the CDC, examples of which are given in the Research Activities section of this document. Several of the Department's adjunct faculty hold appointments in various offices and centers at the CDC, including the Epidemiology Program Office, Center for HIV/AIDS, Center for Environmental Health, Center for Infectious Disease, and Center for Prevention Services.

Emory Winship Cancer Institute

The Emory Winship Cancer Institute is a comprehensive cancer treatment, research and medical training facility recognized nationally and internationally for its capabilities. The Department of Biostatistics serves as a collaborative WCI partner in the areas of biostatistical research and informatics. The mission of the Biostatistics Research and Informatics (BRI) Core is to offer comprehensive, multi-disciplinary resources for the design and conduct of populational, clinical and basic science studies. These include the development of innovative statistical methodology, storage and retrieval of data generated, appropriate statistical analysis, and summarization of the results.

Within this context the Biostatistics and Bioinformatics Research and Informatics Core coordinates and manages statistical activities in the WCI to ensure that investigators have ready access to statistical consultation and support and provides statistical expertise in the design of experiments and studies, including research proposal development, sample size determination, randomization procedures, and plans for interim reviews and final analysis. In collaboration with the Clinical Translational Review Committee, the BRI Core reviews the integrity and statistical soundness of all studies involving human subjects, and interacts with the Clinical Trials and Translational Research Office in the development of protocols and the monitoring and reporting the clinical data.

Biostatistics Collaboration Core

The Biostatistics Collaboration Core (BCC) offers comprehensive statistical consultation and computational services to faculty, staff, and students in the Rollins School of Public Health, the Woodruff Health Sciences Center, and throughout Emory University. Its primary interest is in assuring appropriate use of statistical methodology in all stages of research including preparation of grants and contracts, assistance in analyzing and presenting research data, and statistical review of manuscripts in the publication process.

THE COMMUNITY

Emory University

The 631 acre Emory campus is located in a historic Atlanta suburb about 15 minutes from downtown. The tree-covered, rolling hills provide an ideal college setting for the 11,000 students and the university faculty. The Quadrangle forms the center of the campus and includes many of the university's oldest buildings, several of which are listed on the National Register of Historic Places. Located on the Quadrangle are the Carlos Museum and the Pitts Theology Library, site of the second largest collection of theological titles in North America. Emory has six other libraries for a total of more than two million volumes.

The George W. Woodruff Physical Education Center, a twenty-one-million-dollar recreational complex, is one of the finer sports facilities in the country. It includes seventeen tennis courts, a fifty-meter swimming pool, gymnastics facilities, basketball courts, racquetball and squash courts, indoor and outdoor tracks, a soccer field, badminton and volleyball courts, a dance studio, combatives room, weight room, and human performance laboratory.

Adjacent to the campus is Lullwater, home of the university president. This park-like area of 185 acres includes a 12-acre lake and wooded jogging trails. Lullwater is open to the entire university community.

The Rollins School of Public Health is part of the Robert W. Woodruff Health Sciences Center which also includes the School of Medicine, the Nell Hodgson Woodruff School of Nursing, Emory University Hospital, Emory Midtown Hospital, and the Yerkes Regional Primate Research Center.

The Carter Center of Emory University addresses national and international issues of public policy and brings to Emory a wide range of scholars, government leaders, business executives and other professionals.

Emory University has been invited to join the Association of American Universities, which is made up of 63 of the most productive and accomplished research universities in the country. Emory has also been named to the top category of Research I Universities by the Carnegie Foundation. *U.S. News & World Report* ranked Emory 7th in the nation for its Master of Public Health and 21st overall among national universities. Emory is also currently ranked among the top ten colleges and universities with the largest endowments, with the highest rate of growth in research dollars and grants among universities in the country.

Atlanta

Atlanta is a progressive, dynamic city that has evolved into the cultural, educational, and financial hub of the South. The population of almost 6 million is a mixture of people from the South and other parts of the US along with people from almost every foreign country imaginable, giving the city a rich ethnic and cultural diversity. An efficient rapid rail and city bus system and an advanced highway grid make the entire metro area accessible by both public and private transportation.

Almost any activity is available in the metro area. Atlanta is home to professional baseball, football, basketball, soccer teams, and NASCAR racing. Cultural activities include the Atlanta Symphony Orchestra, the Alliance and Fox Theaters, the Atlanta Ballet, the Atlanta Opera, and the High Museum of Art. Atlanta hosted the 1996 Summer Olympic games.

Atlanta has a dynamic business climate. More than 450 of the Fortune 500 companies have operations in Atlanta. Atlanta has the busiest airport in the US, Hartsfield Jackson International Airport.

Atlanta's climate makes it an extremely livable city, with four distinct seasons. Warm weather extends through the middle of fall; the winter is usually mild with a few snow days each year. The city is nationally known for its brilliant azaleas, rhododendrons, and dogwoods that bloom every spring. Atlanta is just a few hours from the mountains of North Georgia and the Carolinas and from the beaches of Georgia and Florida.

MPH/MSPH DEGREES OFFERED BY THE DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

- Master of Public Health in Biostatistics (MPH)
- Master of Science in Public Health in Biostatistics (MSPH)
- Bachelors of Arts/Bachelors of Science -Master of Science in Public Health in Biostatistics (BA/BS-MSPH)

Which Degree Program Should I Choose?
The degree programs are distinct and it is important to understand which one best fits your needs. The table below details some of the major differences between the programs.

	MPH in Biostatistics	**MSPH in Biostatistics**
Program Focus	Applied Biostatistics, general public health applications	Preparation for doctoral programs in biostatistics, clinical and biomedical statistics, including public health
Credits Required	42	48
Elective Credits	4	9
Thesis Required?	Thesis or Capstone Option	Thesis or Capstone Option
Cohort size	35	25
Sample key skills	Study design, data management, data analysis, survey analysis, in public health	Study design, data analysis, statistical theory in biomedical research (including public health)
Sample Career path	Public health agency, local health department, public health-based NGO, medical school, doctoral study in public health discipline (e.g., epidemiology, environmental health, global health)	Pharmaceutical company, clinical research organization, public health agency, doctoral study in Statistics or Biostatistics

MPH in Biostatistics

The MPH programs in biostatistics offered through the Rollins School of Public Health are designed for individuals with good quantitative skills and interests in the health sciences. Besides courses in biostatistics and epidemiology, the MPH programs include core courses from other public health disciplines. The MPH program provides some training in biostatistics to students who are interested in broader public health related issues.

Students can enter the MPH programs from a variety of academic and professional backgrounds. Some applicants pursue a degree in biostatistics directly after completing undergraduate studies. For others, study in biostatistics is undertaken after completion of medical or public health training or experience. To the extent possible, the curriculum of each student is tailored to his or her background and interests. Students with prior, relevant course work may receive academic credit toward their degree program in biostatistics.

Admission Requirements

A baccalaureate degree from an accredited college or university is required for admission into the MPH program. **Successful completion of multivariate calculus (Calculus III) with a grade of B+ or better are required for admission to the Biostatistics MPH program.** Applicants must submit GRE scores unless they have a doctoral degree in a comparable field. Scores for all sections of the GRE must be above the 50th percentile with a GPA of 3.5 or higher. International applicants whose native language is not English must take the Test of English as a Foreign Language (TOEFL) with a score of 95 or higher.

Applicants to biostatistics are selected on the basis of their quantitative skills and their potential to make a contribution to the practice of biostatistics in a public health setting. Admissions criteria are:

1. Previous studies and grades, especially in quantitative courses such as mathematics, statistics, and computer sciences;

2. Graduate Record Examination (GRE) scores, especially the quantitative and analytic portions;

3. Letters of recommendation that allow the evaluation of the applicant's quantitative abilities and background in public health; and

4. Course work, experience, or interest in health-related subjects.

Tuition and Financial Aid

A small number of merit-based scholarships are available for MSPH and MPH students. US citizens and permanent residents may apply for needs-based financial aid through the Emory Office of Financial Aid. Research assistantships and internships may be available to students in these programs.

Competencies in MPH in Biostatistics

Upon completion of the MPH degree, the graduate will be able to:

- Identify statistical issues in contemporary public health problems.
- Perform power and sample size calculations to assist in the design of clinical or observational studies.
- Use statistical software for advanced data management.
- Analyze continuous data using linear regression models and discrete data using generalized linear models.
- Analyze right-censored data with time-to-event regression models.
- Analyze correlated data (longitudinal and multi-level) using mixed effect and marginal models.
- Explain fundamental concepts of probability and inference used in statistical methodology.

MPH in Biostatistics Course Plan

This is a MPH program for a student who arrives with the required mathematics background.

Year 1 – Fall Semester

Course No.	Title	Credits
BIOS 506	Foundations of Biostatistical Methods	4
BIOS 510	Introduction to Probability Theory	4
BIOS 531	SAS Programming	2
BIOS 590R	Seminar in Biostatistics	1
EPI 530	Epi Methods I	4
PUBH 500	Introduction to Public Health	0

Year 1 – Spring Semester

Course No.	Title	Credits
BIOS 507	Applied Regression Analysis	4
BIOS 511	Intro. to Statistical Inference	4
BSHE 500	Behavioral Sciences in Public Health	2
HPM 500	Intro to US Health Care System	2
BIOS 590R	Seminar in Biostatistics	1
PUBH 501	Interprofessional Education & PH Leadership	0

Year 1 – Summer Semester

Course No.	Title	Credits
BIOS 595	Applied Practical Experience (APE)	0

Year 2 – Fall Semester

Course No.	Title	Credits
BIOS 521	Applied Survival Analysis	2
BIOS 525	Longitudinal & Multilevel Data Analysis	2
BIOS 580	Statistical Practice I	2
GH 500	Critical Issues in Global Health	2
EH 500	Perspectives in Environmental Health	2
BIOS 590R	Seminar in Biostatistics	1

Year 2 – Spring Semester

Course No.	Title	Credits
BIOS 581	Statistical Practice II (Capstone)	2
OR		
BIOS 599R	Thesis	2
BIOS 590R	Seminar in Biostatistics	
	Elective(s)	4

TOTAL CREDIT HOURS: 42

Required Courses: In the MPH program, there are 7 required Biostatistics courses (BIOS 506, BIOS 507, BIOS 510, BIOS 511, BIOS 521, BIOS 525 and BIOS 531) which serve as the foundation of the degree program. Students must attain at least a B- or better in BIOS 506, BIOS 507, BIOS 510, and BIOS 511 to progress to the next level of coursework.

If you earn a letter grade of an F in any BIOS core course than you will have the following options: 1) to take a "Leave of Absence" for the spring semester and return the following Fall semester and retake the failed BIOS core course 2) take elective courses and Public Health Core Courses in the spring semester, or 3) take deep consideration of your career path and is Biostatistics the best fit for me.

Core Courses: A student must take a core course from each of the following departments: Behavioral Sciences & Health Education, Epidemiology, Environmental Health, Health Policy & Management, and Global Health.

Professional Development Courses: Students will register for two required Professional Development courses in their 1st Fall semester, PUBH 500 and PUBH 501. In the spring semester of their 1st year, students will register for PUBH 502. These classes involve mostly self-administered, self-paced activities outside of the classroom with very minimal in-class requirements to enhance professional development, knowledge, and understanding.

Electives: A student has to take from four credit hours of elective courses.

The total number of credit hours required for the degree is 42. To receive the MPH degree, the student must pass all the required and elective courses, including the practicum experience described above, maintain a cumulative GPA of at least B-, and submit either a Capstone or acceptable MPH thesis.

Applied Practice Experience (APE)

An Applied Practice Experience (APE) is a unique opportunity for graduate students to integrate and apply practical skills and training learned through course work and prior experiences in a professional public health environment. All RSPH graduate students are required to complete an APE of at least 200 hours in a public health agency, institution, or community under the supervision of site supervisor and the guidance of the Department. Although there are no credits associated with the APE requirement, the completion of the requirement is noted on the student's transcript.

MSPH in Biostatistics

The MSPH program in biostatistics offered through the Rollins School of Public Health are designed for individuals with good quantitative skills and interests in the health sciences. Besides courses in biostatistics and epidemiology, the MSPH programs include core courses from other public health disciplines. The MSPH program provides rigorous training in biostatistical methods and their applications. Students in this program are trained for positions in government and private health agencies, industry, and research institutes. The MSPH program may also serve as a preparation for entering a doctoral program in biostatistics.

Students can enter the MSPH program from a variety of academic and professional backgrounds. Some applicants pursue a degree in biostatistics directly after completing undergraduate studies. For others, study in biostatistics is undertaken after completion of medical or public health training or experience. To the extent possible, the curriculum of each student is tailored to his or her background and interests. Students with prior, relevant course work may receive academic credit toward their degree program in biostatistics.

Admission Requirements

A baccalaureate degree from an accredited college or university is required for admission into the MSPH program. **Successful completion of multivariate calculus (Calculus III) and a course in linear algebra with a grade of B+ or better are required for admission to the Biostatistics MPH program.** Applicants must submit GRE scores unless they have a doctoral degree in a comparable field. Scores for all sections of the GRE must be above the 50th percentile with a GPA of 3.5 or higher. International applicants whose native language is not English must take the Test of English as a Foreign Language (TOEFL) with a score of 95 or higher.

Applicants to biostatistics are selected on the basis of their quantitative skills and their potential to make a contribution to the practice of biostatistics in a public health setting. Admissions criteria are:

1. Previous studies and grades, especially in quantitative courses such as mathematics, statistics, and computer sciences;

2. Graduate Record Examination (GRE) scores, especially the quantitative and analytic portions;

3. Letters of recommendation that allow the evaluation of the applicant's quantitative abilities and background in public health; and

4. Course work, experience, or interest in health-related subjects.

Tuition and Financial Aid

A small number of merit-based scholarships are available for MSPH and MPH students. US citizens and permanent residents may apply for needs-based financial aid through the Emory Office of Financial Aid. Research assistantships and internships may be available to students in these programs.

Competencies in MSPH in Biostatistics

Upon completion of the MSPH degree the graduate will be able to:

- Identify statistical issues in contemporary public health problems.
- Perform power and sample size calculations to assist in the design of clinical or observational studies.
- Use statistical software for advanced data management.
- Analyze continuous data using linear regression models and discrete data using generalized linear models.
- Analyze right-censored data with time-to-event regression models.
- Analyze correlated data (longitudinal and multi-level) using mixed effect and marginal models.
- Assess the impacts of assumptions in advanced statistical analysis using probability and statistical theory.
- Apply concepts in probability and statistical theory to define performance or extend basic statistical analysis techniques.
- Assess technical accuracy and performance of advanced analytic methods.

MSPH in Biostatistics Course Plan

This is a MSPH program for a student who arrives with the required mathematics background.

Year 1 – Fall Semester

Course No.	Title	Credits
BIOS 508	Biostatistical Methods I	4
BIOS 512	Probability Theory I	4
BIOS 531	SAS Programming	2
BIOS 590R	Seminar in Biostatistics	1
EPI 530	Epi Methods 1	4
PUBH 500	Introduction to Public Health	0

Year 1 – Spring Semester

Course No.	Title	Credits
BIOS 509	Applied Linear Models	4
BIOS 513	Statistical Inference I	4
BSHE 500	Behavioral Sciences in Public Health	2

HPM 500	Intro to US Health Care System	2
BIOS 590R	Seminar in Biostatistics	1
PUBH 501	Interprofessional Education & PH Leadership	0

Year 1 – Summer Semester

Course No.	Title	Credits
BIOS 595	Applied Practical Experience (APE)	0

Year 2 – Fall Semester

Course No.	Title	Credits
BIOS 522	Survival Analysis Methods	2
BIOS 526	Modern Regression Analysis	3
BIOS 580	Statistical Practice I	2
GH 500	Critical Issues in Global Health	2
EH 500	Perspectives in Environmental Health	2
BIOS 590R	Seminar in Biostatistics	1

Year 2 – Spring Semester

Course No.	Title	Credits
BIOS 581	Statistical Practice II (Capstone)	2
OR		
BIOS 599R	Thesis	2
BIOS 590R	Seminar in Biostatistics	1
	Elective(s)	9

TOTAL CREDIT HOURS: 48

Required Courses: In the MSPH program, there are eight required Biostatistics courses (BIOS 508, BIOS 509, BIOS 512, BIOS 513, BIOS 522, BIOS 526, and BIOS 531) which serve as the foundation of the degree program. Students must attain at least a B- or better in BIOS 508, BIOS 509, BIOS 512, and BIOS 513 to progress to the next level of coursework.

If you earn a letter grade of C or lower in BIOS 512 than you cannot go into BIOS 513 in the spring semester. You will have three options: 1) to take a “Leave of Absence” for the spring semester and return the following Fall semester and retake BIOS 512 (this will delay your graduation year for another year- will complete the MSPH in 3 years), 2) take elective courses and Public Health Core Courses in the spring semester and retake BIOS 512 the next fall semester (this will delay your graduation year for another year- will complete the MSPH in 3 years), or 3) transfer to the BIOS MPH degree program and take the required MPH courses in the spring semester.

If you earn a letter grade of an F in any BIOS core course than you will have the following options: 1) to take a "Leave of Absence" for the spring semester and return the following Fall semester and retake the failed BIOS core course (this will delay your graduation year for another year- will complete the MPH or MSPH in 3 years) 2) take elective courses and Public Health Core Courses in the spring semester (this will delay your graduation year for another year- will complete the MSPH in 3 years) or 3) take deep consideration of your career path and is Biostatistics the best fit for me.

Core Courses: A student must take a core course from each of the following departments: Behavioral Sciences & Health Education, Epidemiology, Environmental Health, Health Policy & Management, and Global Health.

Professional Development Courses: Students will register for two required Professional Development courses in their 1st Fall semester, PUBH 500 and PUBH 501. In the spring semester of their 1st year, students will register for PUBH 502. These classes involve mostly self-administered, self-paced activities outside of the classroom with very minimal in-class requirements to enhance professional development, knowledge, and understanding.

Electives: A student has to take nine credit hours of elective courses.

The total number of credit hours required for the degree is 48. To receive the MSPH degree, the student must pass all required core and elective courses, maintain a cumulative GPA of at least B-, and submit an acceptable MSPH thesis.

Applied Practice Experience (APE)

An Applied Practice Experience (APE) is a unique opportunity for graduate students to integrate and apply practical skills and training learned through course work and prior experiences in a professional public health environment. All RSPH graduate students are required to complete an APE of at least 200 hours in a public health agency, institution, or community under the supervision of site supervisor and the guidance of the Department. Although there are no credits associated with the APE requirement, the completion of the requirement is noted on the student's transcript.

BA/BS -MSPH Program in Biostatistics

Emory College and the Rollins School of Public Health (RSPH) offer a 4+1 dual bachelors/master's degree program. Students have an opportunity to complete a Bachelor of Arts (BA) degree or Bachelor of Science (BS) degree in Emory College, most likely with a major concentration in Mathematics, Computer Sciences and Quantitative Sciences, and a Master of Science in Public Health (MSPH) degree in Biostatistics within five years. Emory College students can apply, and may be admitted, to the program during the student's second (sophomore) or third (junior) year. Up to 12 RSPH credit hours taken by Emory College students admitted to the 4+1 program may count towards the required 126 hours of credit for the Bachelor of Arts or Bachelor of Science degree as well as for the required 48 hours for the MSPH degree in Biostatistics. These 12 credit hours are the following: BIOS 508 (4), BIOS 509 (4), and BIOS 513 (4).

Additional RSPH courses taken during enrollment in Emory College may not count towards the BA or BS degree, and students may need to take additional courses in Emory College to meet the graduation requirements for a bachelor's degree. Undergraduate students enrolled in the 4+1 program must enroll in no less than one required undergraduate course during each semester of their junior and senior years.

Admission

Students normally apply for admission to the program early in their third (Junior) year. Applicants to this program will be selected on the basis of performance and promise as Emory undergraduates. The Department of Biostatistics will review student applications (and may choose to interview them as part of the process). Students will learn of their acceptance during either their second (sophomore) or third (junior) year depending upon the track, prior to the time of pre-registration for course in the Fall Semester. In order to strategically plan an academic program that will meet all undergraduate requirements and allow time for the required courses during their fourth (senior) year, students may have to work with their undergraduate academic advisor as early as their second (sophomore) year. It is required that students enrolling in this program will have completed the following undergraduate courses to be considered for admission into the 4+1 degree program: (1) MATH211 (Multivariable Calculus) or MATH 210 (Advanced Calculus for Data Sciences), and (2) MATH 221 (Linear Algebra). Some background and experience in the health sciences is preferable but not required. Students should be in good academic standing at the time of admission.

Advisement

Once students are admitted to the program, they will be assigned an academic advisor in the Department of Biostatistics and Bioinformatics whom they may consult in addition to their undergraduate academic advisor.

Degree Requirements

Emory College, BA/BS: The program does not affect requirements for a major concentration, including College general education and writing requirements. Students will complete the required number of credit hours to graduate with a BA/BS degree at the end of four years. Up to 12 credit hours of RSPH courses will count towards the 126 semester hours required for the BA/BS degree (and also, the required 48 hours for the MSPH degree). While students may enroll in additional RSPH courses while enrolled in an Emory College BA/BS program, no more than 12 credit hours of RSPH coursework may count towards the undergraduate degree, including undergraduate elective requirements.

Tuition and Financial Aid

Students entering the School of Public Health in the 5th year are like any other new student. They are considered dual degree students and are only charged tuition and fees in their 5th year from RSPH. Hence, RSPH only sees them as being at RSPH for one year.

Competencies in BA/BS -MSPH in Biostatistics

Upon completion of the MSPH degree the graduate will be able to:

- Identify statistical issues in contemporary public health problems.
- Perform power and sample size calculations to assist in the design of clinical or observational studies.
- Use statistical software for advanced data management.
- Analyze continuous data using linear regression models and discrete data using generalized linear models.
- Analyze right-censored data with time-to-event regression models.
- Analyze correlated data (longitudinal and multi-level) using mixed effect and marginal models.
- Assess the impacts of assumptions in advanced statistical analysis using probability and statistical theory.
- Apply concepts in probability and statistical theory to define performance or extend basic statistical analysis techniques.
- Assess technical accuracy and performance of advanced analytic methods.

BA/BS -MSPH in Biostatistics Course Plan

There are two tracks for the BIOS 4+1 program.

TRACK A – Default Track: taking BIOS 512 and BIOS 513 in senior year. Students must apply by Feb. 1st of their junior year.

Course No.	Title	Credit
Year 4 (Senior) - Fall Semester		
BIOS 508	Biostatistical Methods	4
BIOS 512	Probability Theory I	4
BIOS 531	SAS Programming	2
BIOS 590R	Seminar in Biostatistics	1
EPI 530	Epi Methods & Lab	4
PUBH 500	Introduction to Public Health	0
Year 4 (Senior) - Spring Semester		
BIOS 509	Applied Linear Models	4
BIOS 513	Statistical Inference I	4
BIOS 590R	Seminar in Biostatistics	1
BSHE 500	Behavioral Sciences in Public Health	2
HPM 500	Intro to US Health Care System	2
PUBH 501	Interprofessional Education & PH Leadership	0
BIOS 595	Applied Practice Experience	0
Year 5 - Fall Semester		
BIOS 522	Survival Analysis Methods	2
BIOS 526	Modern Regression Analysis	3
BIOS 580	Statistical Practice I	2
BIOS 590R	Seminar in Biostatistics	1
GH 500	Critical Issues in Global Health	2
EH 500	Perspectives in Environmental Health	2

Year 5 - Spring Semester

BIOS 581	Statistical Practice II (Capstone)	2
or		
BIOS 599R	Thesis	2
BIOS 590R	Seminar in Biostatistics	1
	Electives	9

TRACK B – taking BIOS 512 and BIOS 513 in junior year. Students must apply by Feb. 1st of their sophomore year.

Course No.	Title	Credit
Year 3 (Junior) - Fall Semester		
BIOS 512	Probability Theory I	4
Year 3 (Junior) - Spring Semester		
BIOS 513	Statistical Inference I	4
Year 4 (Senior) - Fall Semester		
BIOS 508	Biostatistical Methods	4
BIOS 531	SAS Programming	2
BIOS 590R	Seminar in Biostatistics	1
EPI 530	Epi Methods & Lab	4
PUBH 500	Introduction to Public Health	0
Year 4 (Senior) - Spring Semester		
BIOS 509	Applied Linear Models	4
BIOS 590R	Seminar in Biostatistics	1
BSHE 500	Behavioral Sciences in Public Health	2
HPM 500	Intro to US Health Care System	2
PUBH 501	Interprofessional Education & PH Leadership	0
BIOS 595	Applied Practice Experience	0
Year 5 - Fall Semester		
BIOS 522	Survival Analysis Methods	2
BIOS 526	Modern Regression Analysis	3
BIOS 580	Statistical Practice I	2
BIOS 590R	Seminar in Biostatistics	1
GH 500	Critical Issues in Global Health	2
EH 500	Perspectives in Environmental Health	2
Year 5 - Spring Semester		
BIOS 581	Statistical Practice II (Capstone)	2
or		
BIOS 599R	Thesis	2
BIOS 590R	Seminar in Biostatistics	1
	Electives	9

Core Courses: A student must take a core course from each of the following departments: Behavioral Sciences & Health Education, Epidemiology, Environmental Health, Health Policy & Management, and Global Health.

Professional Development Courses: Students will register for required Professional Development courses in their 1st year (senior) at RSPH: Fall semester - PUBH 500 and Spring semester- PUBH 501. These classes involve mostly self-administered, self-paced activities outside of the classroom with very minimal in-class requirements to enhance professional development, knowledge, and understanding.

Electives: A student has to take nine credit hours of elective courses.

The total number of credit hours required for the degree is 48. To receive the MSPH degree, the student must pass all required core and elective courses, maintain a cumulative GPA of at least B-, and submit an acceptable thesis or capstone.

Applied Practice Experience (APE)

An Applied Practice Experience (APE) is a unique opportunity for graduate students to integrate and apply practical skills and training learned through course work and prior experiences in a professional public health environment. All RSPH graduate students are required to complete an APE of at least 200 work hours in a public health agency, institution, or community under the supervision of site supervisor and the guidance of the Department. Although there are no credits associated with the APE requirement, the completion of the requirement is noted on the student's transcript.

BA/BS and MSPH 4+1 students may begin the APE process after no less than 9 credit hours of RSPH coursework is complete. Students may complete the APE the summer between earning the BA/BS degree and beginning the MSPH degree. Students on an F-1 visa, must apply for Optional Practical Training in order to complete the APE requirement in the summer.

COURSE DESCRIPTIONS

BIOS 500 (3) Statistical Methods I (Non-BIOS students): This course covers fundamental concepts and methods used in data analysis. Upon completion of this course, students will understand and be able to apply basic techniques in descriptive and inferential statistics. These include: techniques in graphical and numerical descriptive statistics; elementary probability calculation using the discrete and continuous distributions; point and confidence interval estimation and hypothesis testing for population means and proportions, differences between means and between proportions; elementary nonparametric techniques, and simple linear regression and correlation. Students will use SAS to perform the statistical analyses.

BIOS 500 (1) Lab: The lab portion of BIOS 500 is designed with the main purpose to practice the statistical methods learned in BIOS 500. To achieve this, we teach a few basics of SAS, a widely used statistical software package, as an analytical tool to implement those methods taught in lecture. Please recognize that the purpose of the lab is not to teach complex SAS programming statements or data management.

BIOS 501 (4) Statistical Methods II with Lab (Non-BIOS students)**:** This course is a follow-up to Statistical Methods I, and assumes a certain degree of comfort with these materials. It is composed of three distinct units: (1) covariance, correlation, and simple linear regression, (2) multiple linear regression and model building, and (3) logistic regression and survival analysis. To instill you with confidence in these topic areas, learning will take place through a variety of tasks and assignments administered at frequent intervals, including (but not limited to): problem sets, quizzes, readings, and lab assignments (graded separately). I will use these to evaluate your overall performance, in conjunction with two unit exams and a cumulative final exam. LAB: Learning statistical software is an integral part of learning basic statistical techniques; while many of the ideas and methods we cover are not conceptually difficult, some of the corresponding calculations can be painfully tedious if done by hand. For this course, we chose to use SAS as our statistical software; knowledge of SAS is a common requirement for many jobs in Public Health and related area. By learning SAS, we allow ourselves to not be overly concerned with the mathematical details, and instead can focus on planning appropriate analyses and correctly interpreting the results. Although the lab will have a separate instructor, 25% of your CLASS grade comes from lab, the content is designed to be fully integrated with lecture.

BIOS 502 (3) Statistical Methods III (Non-BIOS students): We start with data analytic methods not covered in BIOS 500 & BIOS 501 (Statistical Methods I & II), such as two-factor and multifactor studies, transformations, time series and forecasting, and missing data and multiple imputation. We then focus on multilevel modeling of intra- and inter-individual change. In addition to building mixed effects models to analyze longitudinal data, more complex hierarchical models will also be formulated to address research questions that involve other types of clustered data. Prerequisites: BIOS 500 and BIOS 501. Because SAS will be used to perform the statistical analyses, students are expected to write SAS programs.

BIOS 506 (4) Foundations of Biostatistical Methods: This course presents a sophisticated introduction to the concepts and methods of biostatistical data analysis. The topics include descriptive statistics; probability; applications of the binomial, Poisson and normal distributions;

sampling distributions; point and confidence interval estimation; hypothesis testing; a variety of one- and two-sample parametric and non-parametric methods for analyzing continuous or discrete data; and simple linear regression. The course will also equip students with computer skills for implementing these statistical methods using statistical software R. Prerequisites: College-level courses in linear algebra and calculus.

BIOS 507 (4) Applied Regression Analysis: This is the first regression analysis course in applied statistics designed for MPH students. Both theoretical and applied aspects of linear regression and generalized linear regression modeling will be covered in this course. The emphasis will be on applications. The first part of the course covers the following topics: simple linear regression, multiple linear regression, analysis of variance, confounding and interaction, residual and influence diagnostics, variable transformations, multicollinearity, model selection and validation. The second part of the course includes: generalized linear models, introduction to maximum likelihood estimation, logistic regression, nominal and ordinal logistic regression, Poisson regression. Parameter interpretation and scientific interpretation of results will be emphasized throughout the course. Students are expected to use SAS (or R), when necessary, for homework assignments.

BIOS 508 (4) Biostatistical Methods: This course is a required course for BIOS MSPH students. It is taken by the BIOS MSPH students and PhD students in the Fall semester of their first year in the program. The prerequisites include College-level courses in Linear algebra and Calculus and programming experience in either SAS or R (or concurrent enrollment in BIOS 531: SAS Programming.) This course provides a mathematically sophisticated introduction to the concepts and methods of biostatistical data analysis. It aims to provide the students the skills to collaborate with investigators and statistical colleagues in the analysis of data from biomedical and public health studies and to communicate the results of statistical analyses to a broad audience. The topics include descriptive statistics; probability; detailed development of the binomial, Poisson and normal distributions and simulation of random variables from these distributions; sampling distributions; point and confidence interval estimation; simulation studies; hypothesis testing; power analysis and sample size calculations; a variety of one- and two-sample parametric and non-parametric methods for analyzing continuous or discrete data and resampling statistics. The course will also equip students with computer skills for implementing these statistical methods using standard statistical software SAS or R.

BIOS 509 (4) Applied Linear Models: The course covers statistical methodology for the analysis of continuous outcome data, primarily from cross-sectional studies and designed experiments. We introduce the key matrix-based methods for estimation and inference based on the multiple linear regression model. Subsequently, topics include secondary hypothesis testing and restrictions, regression diagnostics, model selection, confounding and interaction, analysis of variance and covariance, and an introduction to random effects and the mixed linear model.

BIOS 510 (4) Introduction to Probability TheoryPUBh: Axiomatic probability, random variables, distribution theory, special parametric families of univariate distributions, joint and conditional distributions, distributions of functions of random variables, and probability modeling. Prerequisite: calculus and multivariate analysis.

BIOS 511 (4) Introduction to Statistical Inference : This course provides an introduction to statistical inference. The course is required for Biostatistics MPH students and taken in the second semester of the first year. Fundamental concepts in statistical inference will be covered including: statistical models, sampling distributions, standard errors, asymptotic normality, confidence intervals, hypothesis tests, power analysis. Common frameworks for inference will be discussed including: parametric/likelihood-based inference, the delta method, bootstrap, permutation tests, Bayesian inference.

BIOS 512 (4) Probability Theory I: Introduction to probability, random variables, distributions, conditional distributions, expectations, moment generating functions, and convergence concepts.

BIOS 513 (4) Statistical Inference I: Introduces the theory of parameter estimation, interval estimation, and tests of hypotheses. In this course, we emphasize the classical "frequentist" (i.e., Neyman-Pearson-Wald) approach to inference. As time permits, we briefly explore alternative paradigms of inference such as neo-Fisherian, Bayesian, and statistical decision theory. This course is required for Biostatistics MSPH students and typically is taken in the second semester of the first year.

BIOS 516 (1) Introduction to Large- Scale Biomedical Data Analysis: This is the overview course for the Bioinformatics, Imaging and Genetics (BIG) concentration in the PhD program of the Department of Biostatistics and Bioinformatics. It aims to introduce students to modern high-dimensional biomedical data, including data in bioinformatics and computational biology, biomedical imaging, and statistical genetics. This course will be co-taught by all BIG core faculty members, with each faculty member giving one or two lectures. The focus of the course will be on the data characteristics, opportunities and challenges for statisticians, as well as current developments and hot areas of the research fields of bioinformatics, biomedical imaging and statistical genetics.

BIOS 520 (2) Clinical Trials Methodology: Covers the organization, methodology, and reporting results of clinical trials. Topics covered include: conceptualization, ethical considerations, protocol adherence and compliance, and data collection, as well as statistical techniques such as randomization, double-blind techniques, sample size determination, and analysis considerations. Prerequisites: BIOS 500, or BIOS 506.

BIOS 521 (2) Applied Survival Analysis: This course will provide an introduction to statistical concepts and methods related to the analysis of survival data. Topics include survival functions, hazard rates, types of censoring and truncation, life table, log-rank tests, Cox regression models, and parametric regression models. The emphasis is on practical implementation of standard survival analysis methods using SAS or R and results interpretations.

BIOS 522 (2) Survival Analysis Methods: Deals with the modern methods used to analyze time-to-event data. Background theory is provided, but the emphasis is on using methods and interpreting the results. The course provides coverage of survivorship functions, Kaplan-Meier curves, logrank test, Cox regression, model fitting strategies, model interpretation, stratification, time dependent covariates, and introduction to parametric survival models. Computer programs are used. A data analysis project is required. Prerequisites: BIOS 500 and BIOS 501, or BIOS 506 and BIOS 507.

BIOS 524 (2) Introduction to Analytic Methods for Infectious Diseases: Introduces dynamic and epidemiologic concepts particular to infectious diseases, including the elements of the infection process; transmission patterns, epidemic, endemic, micro- and macroparasitic diseases; zoonoses, basic reproduction number; dependent happenings; and direct and indirect effect of intervention. Prerequisites: Previous or concurrent enrollment in BIOS 500 or BIOS 506, and BIOS 510.

BIOS 525 (2) Longitudinal and Multilevel Data Analysis: This course introduces students to regression techniques commonly used in analyzing longitudinal and multilevel data that are frequently encountered in biomedical and public health research. This course draws motivating examples from environmental and social epidemiology, health services research, clinical studies, and behavioral sciences. The course focuses on data analysis and interpretation. Students will gain practical experience using R/SAS/Stata for statistical computing.

BIOS 526 (3) Modern Regression Analysis: This course introduces students to modern regression techniques commonly used in analyzing public health data. Topics include: (1) parametric and non-parametric methods for modeling non-linear relationships; (2) methods for modeling longitudinal and multilevel data that account for within group correlation; (3) Bayesian regression modeling; and (4) methods for multivariate outcomes. Prerequisites: BIOS 509 or instructor's permission.

BIOS 530 (2) Applied and Advanced Statistics in Observational Studies: This class is designed to cover the concepts and implementations of up-to-date analytic methodologies and strategies in observational studies, and to equip the students with the mindset and essential tools to handle data from observational research either for prediction (statistical learning) or causal inference. Propensity score methods, establishing/validating prediction models, risk stratification, the guidance of Good Research Practice, etc. will be illustrated along with real-life projects and backed up by the recent literatures. The students will also experience a complete research experience from hypothesis development to final analytic report. It is assumed the students finished basic statistical courses, including linear regression models, e.g., BIOS 500 – BIO507, and had the basic skill of using SAS and/or R. Multivariable modeling using logistic regression and Cox proportional hazard regression will be briefly introduced if the students have not been exposed yet. 2nd–Year Master students or Ph.D. students will be a good fit. The 1st-Yr master students should contact the instructor for approval.

BIOS 531 (2) SAS Programming: This course offers instruction in basic SAS programming. It assumes no prior knowledge of SAS, and begins with an introduction to the data step and procedure call. Topics covered include: dataset manipulation, report writing, arrays, looping, simulation, SAS macro, SAS Interactive Matrix Language (IML), SAS Graphics, and SAS Output Delivery System (ODS). The course prepares students to take the Base SAS Certification exam. Students who pass this exam successfully receive a certificate of completion from the SAS Institute. Prerequisites: BIOS 501 or equivalent, OR BIOS 506 (concurrent), OR permission of the instructor.

BIOS 532 (2) Statistical Computing: Programming style and efficiency, data management and data structures, hardware and software, maximum likelihood estimation, matrix methods and least squares, Monte Carlo simulation, pseudo-random number generation, bootstrap, and UNIX-based computing and graphical methods. Prerequisite: BIOS 531, BIOS 506, and BIOS 510, or permission of instructor.

BIOS 534 (2) Machine Learning: This course covers fundamental machine learning theory and techniques. The topics include basic theory, classification methods, model generalization, clustering, and dimension reduction. The material will be conveyed by a series of lectures, homeworks, and projects.

BIOS 540 (2) Introduction to Bioinformatics: This course is an introduction to the field of Bioinformatics for students with a quantitative background. The course covers biological sequence analysis, introductions to genomics, transcriptomics, proteomics and metabolomics, as well as some basic data analysis methods associated with the high-throughput data. In addition, the course introduces concepts such as curse of dimensionality, multiple testing and false discovery rate, and basic concepts of networks.

BIOS 544 (2) Introduction to R programming for Non-BIOS students: The goal of the course is to will provide an introduction to R in organizing, analyzing, and visualizing data. Once you've completed this course you'll be able to enter, save, retrieve, summarize, display and analyze data.

BIOS 545 (2) R Programming for BIOS students: This course covers the basic contents of R programming with applications on statistical data analysis. Topics include data types, language syntax, basic graphics, debugging, creating packages and documentation.

BIOS 550 (2) Sampling Applications: How to select probability samples and analyze data using simple random sampling, stratified random sampling, cluster sampling and multistage sampling. The software package PC-SUDAAN is used for data analysis. Prerequisite: BIOS 500 and 501 or BIOS 506.

BIOS 555 (2) High-Throughput Data Analysis using R and BioConductor: This course covers the basics of microarray and second-generation sequencing data analysis using R/BioConductor and other open source software. Topics include gene expression microarray, RNA-seq, ChIP-seq and general DNA sequence analyses. We will introduce technologies, data characteristics, statistical challenges, existing methods and potential research topics. Students will also learn to use proper Bioconductor packages and other open source software to analyze different types of data and deliver biologically interpretable results.

BIOS 560R (VC) Current Topics in Biostatistics: A faculty member offers a new course on a current topic of interest for both PhD and Master's students.

BIOS 570 (2) Introduction to Statistical Genetics: This is an introductory course for graduate students in Biostatistics, Bioinformatics, Epidemiology, Genetics, Computational Biology, and other related quantitative disciplines. The course will cover statistical methods for the analysis of family and population based genetic data. Topics covered will include classical linkage analysis, population-based and family-based association analysis, haplotype analysis, genome-wide association studies, basic principles in population genetics, imputation-based analysis, pathway-based analysis, admixture mapping, analysis of copy number variations, and analysis of massively parallel sequencing data. Students will be exposed to the latest statistical methodology and computational tools on gene mapping for complex human diseases.

BIOS 580 (2) Statistical Practice I: This course will cover topics dedicated to preparing students to collaborate as biostatisticians for public health and biomedical projects with non-statisticians. Covered topics will include consulting versus collaboration, ethics, nonstatistical aspects of collaboration (e.g. interpersonal communication), and negotiating expectations with clients. The students will work together in small groups to develop research questions based on an existing real life datasets and discussion with clinical collaborator, conduct power analyses, choose the appropriate statistical methodology to analyze the research questions, then answer at least one of the questions, and present the results in both oral and written format. In addition, individually each student will complete a series of milestones that results in oral and/or written proposal for individual capstone project to be completed in the Spring semester.

BIOS 581 (2) Statistical Practice II (Capstone): This is a required course for the MPH and MSPH students in the Biostatistics and Bioinformatics program in their final spring semester. The purpose of the course is to help students with their capstone project in project management, manuscript writing, and oral presentation while they conduct their project with their individual BIOS advisors. Students will review journal articles to critique study design and statistical analysis methods in a journal club format. They will learn how to write journal articles section by section through lectures and homework assignments. They will develop a manuscript based on their capstone project. At the end of the semester, each student will give an oral presentation on his/her capstone project. Each student will also make a poster on his/her capstone project. Students will receive feedbacks from their peers and instructors to improve their writing and presentation skills. The prerequisite is BIOS 580 - Statistical Practice I.

BIOS 584 (3) Python programming for non-BIOS students: This course will provide a pragmatic and hands-on introduction to the Python programming language, with a focus on practical applications and projects, rather than theoretical topics. This course will cover basic programming concepts, data structures, control flow structures, object-oriented programming, and graphical user interface-driven applications. As the course progresses, students will learn to work with Python library packages (such as NumPy, SciPy, Pandas, Matplotlib, and Scikit-learn), data structures, and tools for data science and cyber security. By the end of the course, students will gain a solid working knowledge of Python modules. Students will feel confident about their ability to work with built-in Python packages, write custom code to load, process, and analyze external datasets, generate evaluation metrics, create, and interpret graphs for data visualization and analytics. The examples and problems used in this course will be drawn from diverse areas such as text processing, simple graphics creation, image manipulation, genomics, and electronic medical records (EMR) databases (e.g., Emory Healthcare data sources). Prerequisite: BIOS 500 or equivalent.

BIOS 585 (3) Python Programming for BIOS students: This course will provide a pragmatic and hands-on introduction to the Python programming language, with a focus on practical applications and projects, rather than theoretical topics. This course will cover basic programming concepts, data structures, control flow structures, object-oriented programming, and graphical user interface-driven applications. As the course progresses, students will learn to work with Python library packages (such as NumPy, SciPy, Pandas, Matplotlib, and Scikit-learn), data structures, and tools for data science and cyber security. By the end of the course, students will gain a solid working knowledge of Python modules. Students will feel confident about their ability to work with built-in Python packages, write custom code to load, process, and analyze external datasets, generate evaluation metrics, create, and interpret graphs for data visualization and analytics. The examples and problems used in this course will be drawn from diverse areas such as text processing, simple graphics creation, image manipulation, genomics, and electronic medical records (EMR) databases (e.g., Emory Healthcare data sources). Only BIOS Students.

BIOS 590R (1) Seminar in Biostatistics: features invited speakers, departmental faculty, students, and others who discuss special topics and new research findings.

BIOS 595R (0) Practicum: Enables students to apply skills and knowledge in an applied setting through a supervised field training experience in a public health setting that complements the student's interests and career goals. Must meet RSPH guidelines and have departmental approval.

BIOS 597R (VC) Directed Study: Provides an in-depth exposure to specific topics not covered in regular courses, for example, statistical genetics and specialized experimental designs.

BIOS 599R (VC) Thesis: Master's thesis research.

BIOS 707 (4) Advanced Linear Models: Generalized inverse of a matrix; vectors of random variables; multivariate normal distribution; distribution theory for quadratic forms of normal random variable; fitting the general linear models by least squares; design matrix of less than full rank; estimation with linear restrictions; estimable functions; hypothesis testing in linear regression; and simultaneous interval estimation. Prerequisites: BIOS 507, BIOS 511, and a course in matrix algebra.

BIOS 709 (4) Generalized Linear Models: Studies analysis of data using generalized linear models, as well as models with generalized variance structure. Parametric models include exponential families such as normal, binomial, Poisson, and gamma. Iterative reweighted least squares and quasi-likelihood methods are used for estimation of parameters. Methods for examining model assumptions are studied. Generalized estimating equations (GEE) and quadratic estimating equations are introduced for problems where no distributional assumptions are made about the errors except for the structure of the first two moments. Illustrations with data from various basic science, medicine, and public health settings. Prerequisite: BIOS 511 and BIOS 707.

BIOS 710 (4) Probability Theory II: Axioms of probability, univariate and multivariate distributions, convergence of sequences of random variables, Markov chains, random processes, martingales. Prerequisite: BIOS 512 and BIOS 511.

BIOS 711 (4) Statistical Inference II: Examines the fundamental role of the likelihood function in statistical inference, ancillary and sufficient statistics, estimating functions, and asymptotic theory. This course presents conditional, profile and other approximate likelihoods; various ancillary concepts; generalizations of Fisher information in the presence of nuisance parameters; optimality results for estimating functions; and consistency/asymptotic normality of maximum likelihood and estimation function-based estimators. It briefly discusses alternative approaches to inference including Bayesian, Likelihood Principle, and decision theory. Prerequisite: BIOS 710.

BIOS 722 (2) Advanced Survival Analysis: In-depth coverage of theory and methods of survival analysis, including censoring patterns and theory of competing risks, nonparametric inference, estimating cumulative hazard functions, Nelson estimator, parametric models and likelihood methods, special distributions, two sample nonparametric tests for censored data, power considerations and optimal weights, sample size calculations for design purposes, proportional hazards model, partial likelihood, parameter estimation with censored data, time-dependent covariates, stratified Cox model, accelerated failure time regression models, grouped survival analysis, multivariate survival analysis, and frailty models. Prerequisite: BIOS 510, BIOS 511, BIOS 522.

BIOS 723 (4) Stochastic Processes: Provides dual coverage of the theory and methods for dealing with the diversity of problems involving branching processes, random walks, Poisson processes, and birth and death processes, Gibbs sampling, martingale counting processes, hidden Markov chains, inference on semi-Markov chains and chain of events modeling. Applications will be drawn from the biological sciences, including the theory of epidemics, genetics, survival analysis, and models of birth-migration-death, and the design and analysis of HIV vaccine trials. Prerequisites: Matrix algebra and BIOS 710.

BIOS 724 (2) Analytic Methods for Infectious Disease Interventions: Advanced analytic, statistical, and epidemiologic methods particular to infectious diseases including analysis of infectious disease data and evaluation of intervention. Prerequisites: BIOS 511.

BIOS 726 (2) Applied Multivariate Analysis: This course investigates multivariate techniques. The main subject areas covered are inferences about multivariate means, multivariate regression, multivariate analysis of variance (MANOVA) and covariance (MACOVA), principal components, factor analysis, discriminant analysis and classification, and cluster analysis. Appropriate programs such as SAS and S-PLUS will be demonstrated. Prerequisite: BIOS 507 and BIOS 511.

BIOS 731 (2) Advanced Statistical Computing: This course covers the theories and applications of some common statistical computing methods. Topics include Markov chain Monte Carlo (MCMC), hidden Markov model (HMM), Expectation-Maximization (EM) and Minorization-Maximization (MM), and optimization algorithms such as linear and quadratic programming. The class has two main goals for students: (1) learn the general theory and algorithmic procedures of some widely used statistical models; (2) develop fluency in statistical programming skills. The class puts more emphasis on implementation instead of statistical theories. Students will gain computational skills and practical experiences on simulations and statistical modeling. This course requires significant amount of programming. Each set of homework involves the implementation of certain algorithms using high-level programming language (such as Matlab or R).

BIOS 732 (2) Advanced Numerical Methods: The course covers topics in traditional numerical analysis specifically relevant to statistical estimation and inference. The topics covered include numerical linear algebra, the root finding problem (maximum likelihood) methods such as IRLS, Newton-Raphson, and EM algorithm, and Bayesian techniques for marginalization and sampling for use in statistical inference (MCMC methods). Additional topics may include numerical integration and curve fitting. Prerequisites include BIOS 532, BIOS 710 and BIOS 711, or permission of the instructor. BIOS 711 may be taken concurrently.

BIOS 736 (2) Statistical Analysis with Missing and Mismeasured Data: The goal of the course is to introduce the concepts and methods of analysis for missing data. Topics will include methods for distinguishing ignorable and non-ignorable missing data mechanisms, single and multiple imputation, hot-deck imputation. Computer intensive methods will be used. Prerequisites: BIOS 511 and PhD Biostatistics student.

BIOS 737 (2) Spatial Analysis of Public Health Data: This course will familiarize students with statistical methods and underlying theory for the spatial analysis of georeferenced public health data. Topics covered include kriging and spatial point processes. In addition, review recent computational advances for applying these methods. Prerequisites: BIOS 506, BIOS 507, BIOS 510, BIOS 511.

BIOS 738 (2) Bayesian and Empirical Bayes Methods: Includes Bayesian approaches to statistical inference, point and interval estimation using Bayesian and empirical Bayesian methods, representation of beliefs, estimation of the prior distribution, robustness to choice of priors, conjugate analysis, reference analysis, comparison with alternative methods of inference, computational approaches, including Laplace approximation, iterative quadrature, importance sampling, and Markov Chain Monte Carlo (Gibbs sampling). Various applications such as small area estimation, clinical trials and other biomedical applications will be used as examples. Prerequisite: BIOS 511.

BIOS 745R (2) Biostatistical Consulting: This course will cover topics dedicated to preparing doctoral students to lead biostatistical collaborations with non-statisticians in public health, biology, and medicine academic environments. Covered collaboration topics will include consulting versus collaboration, ethics, non-statistical aspects of collaboration (e.g. interpersonal communication), and negotiating expectations with clients. Covered biostatistical topics will include specific aim refinement, appropriate study design for the research question, assessment of feasibility (time and effort) of different statistical methods for the same problem, statistical review of grant proposals including power calculations, and appropriate summarization/presentation of results to non-statistical audiences. Experience is the best way to nurture the critical thinking skills necessary for excellent biostatistical collaboration. Students will be given weekly assignments to further develop skills in each of the topic areas. Assignment tasks will be drawn from completed projects the course instructors have encountered. In addition, each student, under the mentorship of the course instructors or faculty in the Department of Biostatistics and Bioinformatics will engage in a collaboration experience. Each student will collaborate with a clinical investigator and provide biostatistical support to all aspects of their project. True to real-life experiences, types of projects will vary depending on the investigator and their research question of interest.

BIOS 760R (VC) Current Topics in Biostatistics: A faculty member offers a new course on a current topic of interest for PhD students.

BIOS 761 (4) Causal Inference: This course provides a survey of modern topics in causal inference. Fundamental concepts in causal inference will be covered including: counterfactual random variables, assessing identifiability of causal effects, graphical frameworks, Gcomputation, inverse probability of treatment weighting, methods for efficient, doubly (multiply) robust estimation of causal effects, and causal mediation. Where possible, the course emphasizes the use of modern regression (e.g., machine learning) in causal effect estimation and provides an applied introduction to this area is provided as well. Pre-requisites: BIOS 512 and BIOS 513 or permission of instructor

BIOS 770 (3) Statistical Methods in Human Genetics: This is an introductory course for graduate students in Biostatistics and Bioinformatics, Epidemiology, Computational Biology and other GDBBS disciplines. The course will cover statistical methods for the analysis of genetic and genomic data. Topics covered include population genetics, pedigree analysis, population-based and family-based association analysis, genome-wide association studies (GWAS), analysis of copy number and structural variations, next-generation sequencing data (DNA), analysis of -omics data, and analysis of the microbiome. Students will be exposed to the latest methodology and computer tools in statistical genetics and genomics. Lectures on statistical methodology will be complemented by guest lectures from applied investigators who put these methods into action in a variety of ongoing genetic studies of complex human traits. Individual lectures will be a mixture of didactic instruction (via electronic slides), software illustration, and critical reading of seminal papers. Regarding papers, students will lead presentation and discussion of these seminal papers (see section 2 of Evaluation) during the semester. Students are expected (and encouraged) to ask and answer questions throughout the semester (see section 4 of evaluation). For software illustration, students are encouraged to bring a laptop to class to follow along software demonstrations.

BIOS 777 (1) How to Teach Biostatistics: Prepares students for teaching introductory level courses in biostatistics. The topics discussed are: syllabus development, lecturing, encouraging and managing class discussion, evaluating student performance, test and examinations, cheating, the role of the teaching assistant, teacher-student relationships, teaching students with weak quantitative skills, teaching students with diverse backgrounds, teaching health sciences students, teaching medical students, use of audio-visual techniques, and use of computers. Each student is required to teach a certain subject to the other students and the instructor, followed by a discussion of presentation strengths and weaknesses.

BIOS 780R (1) Research Methods in Biostatistics: Spring. The goals of this course are 1) to provide practical skills and knowledge to complete a PhD dissertation in biostatistics and 2) to introduce students to the research of BIOS faculty. Students will become familiar with the process of PhD research in biostatistical methods. Several topics will be covered including reading academic articles, writing tools and techniques, presentation skills, professional ethics, conducting collaborative research, and high performance computing. Lectures will include presentations by faculty giving an overview of their research with the aim of helping students choose a dissertation advisor and research area.

BIOS 795R (VC) Pre-Candidacy Research: Research pertaining to a dissertation and preparing for the proposal.
BIOS 797R (VC) Directed Study: Provides an in-depth exposure to specific topics not covered in regular courses, for example, statistical genetics and specialized experimental designs.

BIOS 799R (VC) Dissertation: Research pertaining to a dissertation and preparing for the defense.

Data Science

DATA 521 (3) Database Development for Public Health: In this course, you'll learn about the basic structure of relational databases and how to read and write simple and complex SQL statements and advanced data manipulation techniques. By the end of this course, you'll have a solid working knowledge of structured query language. You'll feel confident in your ability to write SQL queries to create tables; retrieve data from single or multiple tables; delete, insert, and update data in a database; and gather significant statistics from data stored in a database. This course will teach key concepts of Structured Query Language (SQL), and gain a solid working knowledge of this powerful and universal database programming language.This course provides a comprehensive introduction to the language of relational databases: Structured Query Language (SQL). Topics covered include: Entity-Relationship modeling, the Relational Model, the SQL language: data retrieval statements, data manipulation and data definition statements. Homework will be done using databases running in MySQL which students install on their machines and proc SQL in SAS. Students develop a real-world database project using MySQL during the course. Prerequisite: SAS proficiency (BIOS 500 or equivalent)

DATA 530 (2) Introduction to Geographic Information Systems: The course introduces the use of geographic information systems (GIS) in the analysis of public health data. We develop GIS skills through homework, quizzes, and a case study. Specific skills include map layouts, visualization, and basic GIS operations such as buffering, layering, summarizing, geocoding, digitizing and spatial queries.

DATA 532 (2) Advanced Geographic Information Systems: The course continues the use of geographic information systems (GIS) in the analysis of public health data and adds more advanced features. We develop GIS skills through homework, quizzes and a final project, and particularly build upon the skills learned in INFO 530 such as map layouts, visualization, basic spatial statistics, and basic GIS operations such as buffering, layering, summarizing, geocoding, digitizing and spatial queries. We add new topics such as raster analysis open source GIS, (qgis), geo databases, story maps, and making maps in R. Prerequisites: Experience with Windows-based computing, INFO 530 (or permission of instructor), knowledge of R (BIOS 544 or permission of instructor)

DATA 534 (2) Applied Machine Learning: The elective course gives an introduction to machine learning techniques and theory, with a focus on its use in practical applications. The Applied Machine Learning course teaches you a wide-ranging set of techniques of supervised and unsupervised machine learning approaches using R as the programming language. During the course, a selection of topics will be covered in supervised learning, such as linear models for regression and classification, or nonlinear models such as neural networks, and in unsupervised learning such as clustering. The uses and limitations of these algorithms will be discussed, and their implementation will be investigated in programming assignments. The course also covers theoretical concepts such as inductive bias, the PAC and Mistake-bound learning frameworks, minimum description length principle, and Ockham's Razor. There will be a strong emphasis on the real-world context in which machine learning systems are used. The use of machine learning components in practical applications will be exemplified, and Public health realistic scenarios will be studied in application areas such as hospitalization metrics using electronic medical record data, clinical trials, natural language processing, image processing, and bioinformatics. The importance of the design and selection of features, and their reliability, will be discussed. In order to ground these methods the course includes some programming and involvement in a semester-long research project. This is a programming course: you will be required to write code. Prerequisite: BIOS 500, BIOS 544 (or BIOS 545) or permission of instructor

DATA 550 (2) Data Science Toolkit: This course is an elective for Masters and PhD students interested in learning some fundamental tools used in modern data science. Together, the tools covered in the course will provide the ability to develop fully reproducible pipelines for data analysis, from data processing and cleaning to analysis to result tables and summaries. By the end of the course students will have learned the tools necessary to: develop reproducible workflows collaboratively (using version control based on Git/GitHub), execute these workflows on a local computer (using command line operations, RMarkdown, and GNU Makefiles), execute the workflows in a containerized environment allowing end-to-end reproducibility (using Docker), and execute the workflow in a cloud environment (using Amazon Web Services EC2 and S3 services). Along the way, we will cover a few other tools for data science including best coding practices, basic python, software unit testing, and continuous integration services.

DATA 555 (2) Current Topics in Data Science: This course is the culminating experience of the data science certificate program and is to be taken in the spring semester of second year. The course provides logistical support to students in completing the Data Science Certificate-required integrated learning experience (ILE). The course additionally reviews current topics of interest in data science and helps prepare students for the data science job market. The first several meetings of this course focus on helping students identify suitable data science products and planning for the skills and tools that are needed to complete the ILE-related requirements for the data science certificate. Subsequent classes will cover modern topics in data science (e.g., R Shiny, communicating with diverse audiences, software unit testing, data sharing and privacy) and lectures on preparations for applying for data science-related jobs.Students should be in their second year and be in good standing for completing the data science certificate requirements at the time of enrolling in the class.

DATA 556 (2) Data Visualizations in Public Health: This course is all about data visualization in public health: the art and science of turning data into readable and interpretable graphics. We will explore how to design and create data visualizations based on real world data available and tasks to be achieved. This process includes data modeling, data processing (such as aggregation and filtering), mapping data attributes to graphical attributes, and strategic visual encoding based on known properties of visual perception as well as the task(s) at hand. Students will also learn to evaluate the effectiveness of visualization designs, and think critically about each design decision, such as choice of color and choice of visual encoding. Students will create their own data visualizations including dashboards. Students will also read papers from the current and past visualization literature and create video presentations of their findings. Students will use Tableau as their main tool to visualize data and develop dashboards but will develop transferrable skills which can apply to many of the most popular software packages in the current marketplace. Prerequisite: BIOS 544 or equivalent, or permission from the instructor.
IMPORTANT: This is a hands-on course. You will need to have your computer (Mac or PC) up to date to install, Tableau Desktop software (provided free to students in the class).

Contact Information

Melissa Sherrer, M.Ed
Senior Associate Director of Academic Programs
Department of Biostatistics and Bioinformatics
Rollins School of Public Health
Emory University
1518 Clifton Road
GCR, Room 316
Atlanta, GA 30322
Office: 404-727-3968
msherre@emory.edu