

THE DEPARTMENT OF
BIOSTATISTICS
STUDENT HANDBOOK

2020–2021

Message from the Chair



Welcome to the Department of Biostatistics at Columbia University's Mailman School of Public Health! There has never been a more exciting time to study biostatistics. The advent of data science, the increasing availability of big data, and stark reminders of

public health's importance such as the COVID-19 pandemic that has engulfed our lives have made the need for rigorous analytic tools, and their proper use to address scientific hypotheses extremely important. By studying biostatistics, you will be able to draw sound inferences by separating the signal from the noise and affect proper interpretations that would advance public health for all.

By joining our highly reputable graduate programs, you will be trained in the rigorous theory that underlies statistical principles and the latest computational tools and machine learning techniques to prepare you for the age of data science. You will get ample opportunity to immerse in biomedical problems from complex networks that would inspire methodologic advances and allow for comprehensive data analysis with proper attention to validity of assumptions. Members of our faculty are at the forefront of development of cutting-edge techniques to address important public health challenges, making significant impact in numerous areas including multi-omics analysis, personalized medicine, mobile health, analysis of electronic medical records, and clinical trials.

Your experience will be enriched by interactions with fellow students from diverse backgrounds, providing excellent training in theory and methods, as well as skills necessary for professional practice. As our alumni would attest, employers recognize the high quality of our training programs, and our graduates receive premier jobs in academia, private sector, and government.

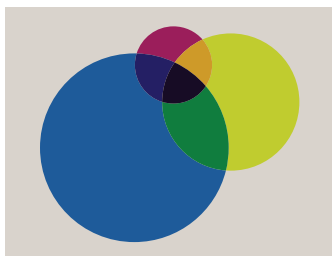
We fully realize that you are joining us at a time of great uncertainty due to the COVID-19 pandemic. I would like to assure you that the Department of Biostatistics is fully committed to being available and accessible to ensure the full Columbia experience by providing all necessary help and guidance. In this handbook, you will find readily accessible information on faculty research interests, descriptions of our state-of-the-art course offerings and coursework, and milestones for our programs.

Welcome again as you join our talented faculty, staff, students, and alumni. I look forward to following your academic and professional growth with great interest. I am always happy and ready to hear from you and to help make your stay with us as fulfilling as possible.

A handwritten signature in black ink, reading "Kiros B." with a stylized flourish.

Kiros Berhane, PhD

Cynthia and Robert Citron-Roslyn and Leslie Goldstein
Professor and Chair
Department of Biostatistics



This visual representation of biostatistics allows us to describe our discipline through various intersections and overlaps. Biostatistics can be thought of as being at the intersection of science, art, and discovery. It can be thought of as the catalyst among interdisciplinary team members working together to solve a problem in public health or medicine. It also allows us to recall that systems are complex, and that our design and analytic approaches need to be developed within that context. The issues are usually more complex than we initially realize.

The Department of Biostatistics

Biostatistics is the science of developing and applying statistical methods for quantitative studies in biomedicine, health, and population sciences.

Biostatisticians play a crucial role in research design, collection and organization of data, analysis, presentation, and interpretation of results. Career opportunities are usually found in governmental agencies, private industry, and medical research institutions.

The Department of Biostatistics maintains collaborative relationships with other units of the University and with outside agencies and institutions. Among the many affiliated institutions and centers are: Columbia University Medical Center, New York State Psychiatric Institute, the Department of Statistics at Columbia's Morningside Campus, the Gertrude H. Sergievsky Center (research in the field of neuroepidemiology), the Herbert Irving Comprehensive Cancer Center and Institute of Cancer Research, the HIV Center for Clinical and Behavioral Studies, and the Irving Center for Clinical Research.

Faculty in the Department of Biostatistics work at the frontier of public health, leading research teams that investigate some of today's most pressing health issues. Recruited from the top universities from around the world, the faculty bring to the School a wealth of experience that serves to inform their research and teaching.

HOWARD ANDREWS (PhD, Rutgers University)

Associate Professor at CUMC of Neuroscience (in Biostatistics)

Research interests: Clinical trials, data management systems, multi-level analysis, perinatal outcomes, environmental factors, Alzheimer's disease

KIROS BERHANE (PhD, University of Toronto)

Cynthia and Robert Citron-Roslyn and Leslie Goldstein Professor and Chair

Research interests: Longitudinal data modeling, multi-level growth curve models, nonparametric regression, multiple outcomes, quantile regression, mediation, applications to environmental data

MELISSA D. BEGG (ScD, Harvard University)

Professor of Clinical Biostatistics, and Dean of the Columbia School of Social Work

Research interests: Analysis of clustered data, oral health research, mental health statistics, clinical research training

QIXUAN CHEN (PhD, University of Michigan)

Associate Professor of Biostatistics

Research interests: Bayesian inference for complex survey data, analysis of incomplete data, non-parametric regression, and random effects models

BIN CHENG (PhD, University of Wisconsin-Madison)

Professor of Biostatistics

Research interests: Linear and generalized linear mixed models, statistical analysis of clinical trials, longitudinal non-normal data modeling, statistical computing, statistical inference on manifolds

YING-KUEN KENNETH CHEUNG (PhD, University of Wisconsin-Madison)

Professor of Biostatistics

Research interests: Design and analysis of clinical trials, methods in toxicology studies and bioassay, applications of Monte Carlo methods, nonparametric methods, bioethics

CODRUTA CHIUZAN (PhD, Medical University of South Carolina)

Assistant Professor of Biostatistics

Research interests: Applications of statistical methods in psychiatric research, functional data analysis, statistical machine learning, treatment regime estimation and evaluation, integration of multi-source data

DEBRA D'ANGELO (MS, Columbia University)

Associate in Biostatistics

Research interests: Applied biostatistical consulting in various medical specialties, database development and data management for research studies, SQL programming

HANGA GALFALVY (PhD, University of Illinois)

Associate Professor of Biostatistics (in Psychiatry)

Research interests: Statistical methodology in psychiatric research, with a special focus on the prediction models for suicidal behavior from high-dimensional data, censored regression models, statistical genetics, and longitudinal data analysis in observational studies

JEFF GOLDSMITH (PhD, Johns Hopkins University)

Associate Professor of Biostatistics

Research interests: Functional data analysis, high-dimensional regression, longitudinal data analysis, smoothing, Bayesian variable selection, neuroimaging, and accelerometry

PRAKASH GORROOCHURN (PhD, Monash University)

Associate Professor of Clinical Biostatistics

Research interests: Mathematical population genetics, genetic mapping of complex diseases

JIANHUA HU (PhD, University of North Carolina-Chapel Hill)

Professor of Biostatistics (in Medicine)

Research interests: high-dimensional genomics/proteomics, imaging, and longitudinal data, modeling disease heterogeneity, and adaptive designs to achieve personalized treatments

IULIANA IONITA-LAZA (PhD, New York University)

Associate Professor of Biostatistics

Research interests: Statistical genetics and bioinformatics

HAOMIAO JIA (PhD, Case Western University)

Professor of Biostatistics (in Nursing)

Research interests: Small area estimation, data smoothing, temporal-spatial analysis, survey sampling

ZHEZHEN JIN (PhD, Columbia University)

Professor of Biostatistics

Research interests: Survival analysis, resampling methods, ROC curves, smoothing methods, nonparametric regression, clinical trials

SEONJOO LEE (PhD, University of North Carolina-Chapel Hill)

Assistant Professor of Clinical Biostatistics (in Psychiatry)

Research interests: Neuroimaging, cognitive neuroscience, machine learning, and functional data analysis

SHING M. LEE (PhD, Columbia University)
Associate Professor of Clinical Biostatistics

Research interests: Rapid dose finding techniques in Phase I trials, and the development of more sensitive endpoints (e.g. Toxicity Burden Scores) in Phase I Trials

CHENG-SHIUN LEU (PhD, Columbia University)
Professor of Clinical Biostatistics (in Psychiatry)

Research interests: Sequential selection procedures for multi-armed clinical trials, statistical application in behavioral studies

CHRISTINE MAURO (PhD, Columbia University)
Assistant Professor of Biostatistics

Research interests: Analysis of clinical trials, longitudinal data analysis, statistical learning techniques, and the application of statistics to problems in mental health research

IAN MCKEAGUE (PhD, University of North Carolina at Chapel Hill)
Professor of Biostatistics

Research interests: Survival analysis, competing risks in HIV/AIDS studies, inference for stochastic processes, empirical likelihood, Markov chain Monte Carlo, functional data analysis, semiparametric efficiency, Bayesian statistics, and martingale and counting process methods

DANIEL MALINSKY (PhD, Carnegie Mellon University)
Assistant Professor of Biostatistics

Research interests: Causal inference, graphical models, missing data, stochastic processes, machine learning, algorithmic fairness, social & environmental determinants of health, health disparities

CALEB MILES (PhD, Harvard University)
Assistant Professor of Biostatistics

Research interests: Causal inference, HIV, Interference, Measurement error, Mediation analysis, Semiparametric inference

TODD OGDEN (PhD, Texas A&M University)
Professor of Biostatistics (in Psychiatry)

Research interests: Analysis of brain imaging data, functional data analysis, nonparametric regression, wavelet applications, statistical modeling

MARTINA PAVLICOVA (PhD, Ohio State University)
Associate Professor of Clinical Biostatistics

Research interests: Functional magnetic resonance imaging, multiple comparisons methods, spatial statistics

MIN QIAN (PhD, University of Michigan)
Associate Professor of Biostatistics

Research interests: Medical decision making, dynamic treatment regimes, variable selection/model selection for decision making, statistical machine learning, reinforcement learning, statistical inference, bootstrap, empirical processes, concentration inequalities, stochastic processes

YIFEI SUN (PhD, Johns Hopkins University)
Assistant Professor of Biostatistics

Research interests: General biostatistical methodology for survival, longitudinal and multivariate data, machine learning, electronic health record data, wearable device data

JOHN L.P. (SEAMUS) THOMPSON (PhD, University of California-Los Angeles)
Clinical Professor of Biostatistics and Neurology

Research interests: Randomized clinical trials, trial design, neurology, data management systems

NAITEE TING (PhD, Colorado State University)
Adjunct Professor

Research interests: Clinical development of new drugs, dose selection, Phase II

WEI-YANN TSAI (PhD, University of Wisconsin-Madison)
Professor of Biostatistics

Research interests: Survival analysis, incomplete data methods, nonparametrics

LINDA VALERI (PhD, Harvard University)
Assistant Professor of Biostatistics

Research interests: Causal inference, measurement error, missing data, mental health, environmental health, and health disparities

MELANIE WALL (PhD, Iowa State University)
Professor of Biostatistics (in Psychiatry)

Research interests: Latent variable modeling, spatial, and longitudinal data analysis

SHUANG WANG (PhD, Yale University)
Professor of Biostatistics

Research interests: Statistical genetics, genetic epidemiology, quantitative trait loci analysis

YUANJIA WANG (PhD, Columbia University)
Professor of Biostatistics

Research interests: Statistical learning, analytics for personalized medicine, and network analysis; applications to psychiatric disorders and neurological disorders.

YING WEI (PhD, University of Illinois-Urbana Champaign)
Professor of Biostatistics

Research interests: Quantile regression methods, growth charts estimation, longitudinal data analysis, semiparametric modeling, and robust statistics

PRIYA J. WICKRAMARATNE (PhD, Yale University)
Associate Professor of Clinical Biostatistics

Research interests: Epidemiologic methods, observational studies, survival analysis, generalized linear models, psychiatric epidemiology

Columbia MPH

Director: Martina Pavlicova, PhD

The **Columbia Masters in Public Health in Biostatistics (MPH)** is a two-year program designed to enhance the quantitative skills of public health practitioners who use statistics frequently in their work. This degree is intended primarily for specialists in public health who use and adapt statistical procedures for health and medical care programs, or serve in a technical capacity as resource person and collaborators in field and programmatic studies.

Course Requirements

The structure of the degree program includes five components, which are all carefully timed and integrated so that learning in one part of the program informs activities and assignments in another:

1. **Discipline** - courses required by your home department
2. **Core** - curriculum that provides the broad, interlocking foundation of knowledge needed for a career in public health
3. **Integration of Science and Practice** - course that bridges the gap between what you traditionally learn in a classroom and the real-world experience of working as a public health professional
4. **Leadership & Development** - course that aims to develop and improve MPH students' abilities in three key areas: leading teams in a variety of settings, working effectively as a team member, and implementing fresh, innovative ideas within an organization or larger community.
5. **Practicum and Capstone Paper** - internship in the field with final paper.

Certificate

Every student in the two-year MPH program enrolls in a certificate program which provides training in a focused area of expertise—in addition to the student's departmental discipline—and leads to a Columbia University approved credential. The certificate programs have been developed in consultation with public health employers and other key stakeholders and reflect today's most sought-after skills and knowledge.

Detailed information on all certificate programs is available [online](#). Students taking the Columbia MPH within the Department of Biostatistics are able to select a certificate from various school-wide certificate programs listed below. You can also find requirements and sample coursework for each certificate using the [Certificate Requirements Database](#).

Advanced Epi
Child Youth and Family Health
Climate and Health
CEOR
Environmental Health Policy
Epi of Chronic Disease
Health and Human Rights

Health Communication
Health of an Aging Society
Health Policy and Practice
Health Promotion Research and
Practice
History, Ethics, Law
Infectious Disease

Injury Prevention and Control
Molecular Epi
Public Health Research Methods
Sexuality, Sexual, and Repro Health
Social Determinants of Health

Practicum

One term of practical experience is required of all students, providing educational opportunities that are different from and supplementary to the more academic aspects of the program. The practicum may be fulfilled during the school year or over the summer. MPH students are required to do a minimum of 280 hours (8-10 full-time weeks) in a public health setting.

Prior to beginning a practicum, it is the student's responsibility to complete a **Scope of Work (SOW)**. While the Office of Field Practice maintains centralized forms for recording the students SOW, approval of the SOW occurs at the departmental level. Students must consult with their faculty advisor who must approve both the proposed practicum project prior to its initiation, and the report submitted at the conclusion of the practicum experience.

Students will be required to make a poster presentation at the department's Annual Practicum Poster Symposium which is held in early May. See the **Practicum Information Section** at the back of the handbook for details and links to required forms.

Capstone Experience

A formal culminating experience is required for graduation. The capstone consulting experience is designed to enable students to demonstrate their ability to integrate their academic studies with the role of biostatistical consultant/collaborator, which will comprise the major portion of their future professional practice. Students register for P8185 Capstone Consulting Seminar, a one-semester, one-credit course during their final spring semester. The course requires students to attend at least one session of the Biostatistics Consulting Service and present the consult to the class for discussion. The Consulting Service, which is run by faculty in the Department of Biostatistics, offers advice on data analysis and appropriate methods of data presentation for publications. It also provides design recommendations for public health and clinical research, including preparation of grant proposals and manuscripts. Operating five days a week, the Consulting Service is a free biostatistical resource for public health and other health sciences researchers at Columbia University.

Curriculum

Required Discipline Courses		Points
P6110	Statistical Computing Using SAS	3
P8100	Applied Regression I	3
P8107	Introduction to Mathematical Statistics	3
P8110	Applied Regression II	3
P8120	Analysis of Categorical Data	3
P8185	Capstone Consulting Seminar	1
TOTAL POINTS FROM REQUIRED COURSES		16

Elective Courses		Points
89260	Building Interdisciplinary Research models	2
P8101	Introduction to Health Data Science	3
P8105	Data Science I	3
P8106	Data Science II	3
P8140	Randomized Clinical Trials	3
P8142	Clinical Trial Methodology	3
P8144	Pharmaceutical Statistics	3
P8157	Analysis of Longitudinal Data	3
P8158	Latent Variable and Structural Equation Modeling for Health Sciences	3
P8160	Topics in Advanced Statistical Computing	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3

Timeline

Fall I	Spring I	Fall 2	Spring 2
MPH Core	Integration of Science and Practice	P8107 Introduction to Mathematical Statistics	P8185 Capstone Consulting Seminar
	Leadership and Development	P8110 Applied Regression II	Certificate Requirements
	P6110 Statistical Computing Using SAS	Certificate Requirements	Complete practicum requirements
	P8100 Applied Regression I		
	P8120 Analysis of Categorical Data		

Master of Science degree programs

The Department of Biostatistics offers two Master of Science degree programs: the MS in Biostatistics and the MS in Patient Oriented Research. Students pursuing the MS in Biostatistics degree select one of four tracks of specialization: Clinical Research Methods, Pharmaceutical Statistics, Statistical Genetics, and Theory & Methods. The MS in Patient Oriented Research degree program is also housed in the department. Whether the focus of the degree is to prepare for doctoral research training, to advance the skills critical for clinical scientists, or as a biostatistician in public health or the pharmaceutical industry, both programs require a facility for quantitative reasoning and a true enjoyment of working with data.

Upon satisfactory completion of the MS in Biostatistics or the MS in Patient Oriented Research, graduates will be able to:

Data Analysis and Computing

- Formulate and produce graphical displays of quantitative information (e.g., scatter plots, box plots, line graphs) that effectively communicate analytic findings;
- Explain general principles of study design in attempting to identify risk factors for disease, isolate targets for prevention, and assess the effectiveness of one or more interventions;
- Select and perform appropriate hypothesis tests for comparing two or more independent exposure groups, or two or more groups of matched/clustered subjects, with respect to a discrete or continuous response measurement of interest;
- Interpret associations estimated via linear regression, logistic regression, and Cox models for survival data;
- Apply the basic tenets of research design and analysis for the purpose of critically reviewing research and programs in disciplines outside of biostatistics; and
- Interpret quantitative findings in accurate, accessible language for colleagues outside of biostatistics, as well as for broader dissemination to the public and other public health professionals.

Public Health and Collaborative Research

- Translate research objectives into testable hypotheses;
- Compare and contrast different study designs and their implications for inference in medical/public health research;
- Describe basic principles and the practical importance of key concepts from probability and inference (including random variation, systematic error, sampling error, measurement error, hypothesis testing, type I and type II errors, confounding bias, and effect modification) to colleagues without extensive statistical training;
- Develop and execute power and sample size calculations for research studies utilizing simple random sampling; and
- Evaluate research reports and proposals for research funding on the basis of their scientific integrity, validity, and the strength of the quantitative analysis.



A brief comparison of the MS Degree Programs

Degree Program	Track	Minimum Credits	Typical Duration	Practicum	Capstone
Master of Science in Biostatistics	Clinical Research Methods Track (MS/CRM)	30	4 semesters	No	Yes
Master of Science in Biostatistics	Pharmaceutical Statistics Track (MS/PS)	35	4 semesters	Yes	Yes
Master of Science in Biostatistics	Statistical Genetics Track (MS/SG)	36	4 semesters	Yes	Yes
Master of Science in Biostatistics	Theory and Methods Track (MS/TM)	36	4 semesters	Yes	Yes
Master of Science in Patient Oriented Research	Patient Oriented Research Program (MS-POR)	30	5 semesters (including summer)	No	Yes

Clinical Research Methods Track

Director: Todd Ogden, PhD

The Master of Science in Biostatistics degree program - Clinical Research Methods track (MS/CRM) provides formal, rigorous training in skills critical to the design and analysis of clinically oriented research studies. It is intended for physicians, nurses, dentists, psychologists, pharmacists, and other health care professionals who plan careers or are actively engaged in clinical research. MS/CRM students will hone their quantitative talents to better pursue research objectives in their chosen fields. As the level of competitiveness for limited research support increases, it is now more important than ever to develop a well-designed study with a strong analytic plan. Mastery of applied biostatistical methods improves the likelihood of assembling compelling and effective clinical research projects and promoting good research practices.

Course Requirements

The required courses are intended to enable degree candidates to gain proficiency in study design, application of commonly-used statistical procedures, facility with statistical software packages, and ability to successfully interpret and communicate the results of an analysis. Students must complete a minimum of 30 points to earn the MS/CRM degree, of which 24 points must be taken at the Mailman School of Public Health. *Up to two electives may be taken pass/fail, especially to encourage students to take courses outside their field of experience.*

Note that some courses in the required curriculum may be waived based on prior coursework with approval of the faculty advisor. In this event, the student may substitute another, more advanced course in place of the waived course. Students interested in completing the program in 1.5 years, are strongly advised to begin coursework by enrolling in the **Columbia Summer Research Institute** which allows students to complete 10 credits over 5 weeks.

In advance of beginning the MS program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health. Students' progress will be reviewed after each semester. Those students whose academic performance falls below a B average (3.0 GPA) in required courses may not be allowed to graduate without remedial course work.

Capstone Experience

As part of the MS/CRM training, each student is required to enroll in P9160 Master's Essay—Clinical Research Methods. This research component of the MS/CRM program should be completed during the final year of study. Conducted in workshop style, students in this course will participate in a weekly seminar geared towards enhancing research skills. At the end of the term, each student will be required to submit a research paper of publishable quality, summarizing his or her research project. Students planning on taking P9160 should have a data set of interest available to them; you must ensure that you have permission (and IRB approval) to analyze and publish results from your analysis. Most sessions will be devoted to discussion of the individual research projects and related current literature. Other times, active researchers from biostatistics and various clinical disciplines may be invited to speak on research techniques, statistical methodology as applied in current studies, and their experiences in conducting patient oriented research. Students will present their topics, plans for analysis, and interpretation of their findings to the class for evaluation and feedback. The completion and submission of this research paper satisfies the student's practicum requirement.

Curriculum (TOTAL POINTS: 30 OR MORE)

Required Courses		Points
P6104	Introduction to Biostatistical Methods	3
P6400	Principles of Epidemiology	3
P8100	Applied Regression I	3
P8110	Applied Regression II	3
P8120	Analysis of Categorical Data	3
P8140	Introduction to Randomized Clinical Trials	3
P8438	Epidemiology II: Design and Conduct of Observational Epidemiology	3
P9160	Master's Essay - Clinical Research Methods	3
TOTAL POINTS FROM REQUIRED COURSES		24

Elective Courses		Points
<i>Choose 2 or more courses from this list or from alternatives approved by your academic advisor.</i>		
P6110	Statistical Computing with SAS	3
P6530	Issues and Approaches in Health Policy and Management	3
P8104	Probability	3
P8105	Data Science I	3
P8109	Statistical Inference	3
P8112	Systematic Review and Meta-analysis (offered through CSRI)	1.5
P8142	Clinical Trial Methodology	3
P8144	Pharmaceutical Statistics	3
P8149	Human Population Genetics	3
P8157	Analysis of Longitudinal Data	3
P8158	Latent Variable and Structural Equation Modeling for Health Sciences	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3
P8307	Molecular Epidemiology	3
P8308	Molecular Toxicology	3
P8404	Epidemiology of Genetics and Aging	3
P8405	Genetics in Epidemiology	3
P8406	Epidemiology of Infections Diseases I	3
P8414	Cancer Epidemiology	3
P8417	Selected Problems of Measurement in Epidemiology	3
P8432	Environmental Epidemiology	3
P8440	Epidemiology of Cardiovascular Diseases	3
P8482	Outcomes Research: Methods and Public Health Implications	3
P8545	Analysis of Large Scale Data Sets (offered through CSRI)	1.5

Sample Timeline

Fall I	Spring I	Fall II	Spring II
P6104 Introduction to Biostatistical Methods	P8100 Applied Regression I	P8110 Applied Regression II	P8438 Epidemiology II
P6400 Principles of Epidemiology I	P8120 Analysis of Categorical Data	P8140 Introduction to RCTs	P9160 Master's Essay
	Elective	Elective	

Pharmaceutical Statistics

Director: Ken Cheung, PhD

The Master of Science in Biostatistics degree program - Pharmaceutical Statistics track (MS/PS) provides study design, research, and biostatistics skills to individuals who are currently working in the pharmaceutical research industry and those seeking to begin a career in the industry. MS/PS students will understand the challenges and modern methods relevant to translational research and clinical trials.

Course Requirements

Students must complete a minimum of 35 credits of coursework to earn the MS/PS degree, of which 30 points must be taken at the Mailman School of Public Health. *Up to two electives may be taken pass/fail, especially to encourage students to take courses outside their field of experience.*

Note that some courses in the required curriculum may be waived based on prior coursework with approval of the faculty advisor. In this event, the student may substitute another, more advanced course in place of the waived course.

In advance of beginning the MS program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health. Students' progress will be reviewed after each semester. Those students whose academic performance falls below a cumulative B average (3.0 GPA) in required courses may not be allowed to graduate without remedial course work.

International students are required to be registered for at least 12 credits during their second and third semesters.

Practicum Requirement

One term of practical experience is required of all students, providing educational opportunities that are different from and supplementary to the more academic aspects of the program. The practicum may be fulfilled during the school year or over the summer. Arrangements are made on an individual basis in consultation with faculty advisors who must approve both the proposed practicum project prior to its initiation, and the report submitted at the conclusion of the practicum experience. Students will be required to make a poster presentation at the department's Annual Practicum Poster Symposium which is held in early May. See the [practicum information section](#) at the back of the handbook for details and links to required forms.

Capstone Experience

A formal culminating experience is required for graduation. The capstone consulting experience is designed to enable students to demonstrate their ability to integrate their academic studies with the role of biostatistical consultant/collaborator, which will comprise the major portion of their future professional practice. Students register for P8185 Capstone Consulting Seminar, a one-semester, one-credit course during their final spring semester. The course requires students to attend at least one session of the Biostatistics Consulting Service and present the consult to the class for discussion. The Consulting Service, which is run by faculty in the Department of Biostatistics, offers advice on data analysis and appropriate methods of data presentation for publications, and provides design recommendations for public health and clinical research, including preparation of grant proposals and manuscripts. Operating five days a week, the Consulting Service is a free biostatistical resource for public health and other health sciences researchers at Columbia University.

Curriculum (TOTAL POINTS: 35 OR MORE)

Required Courses		Points
P6110	Statistical Computing with SAS	3
P6170	New Drug Development: A Regulatory Overview	3
P6400	Principles of Epidemiology	3
P8104	Probability	3
P8120	Analysis of Categorical Data	3
P8130	Biostatistical Methods I	3
P8140	Introduction to Randomized Clinical Trials	3
P8142 *	Clinical Trial Methodology	3
P8144	Pharmaceutical Statistics	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3
P8185	Capstone Consulting Seminar	1
TOTAL POINTS FROM REQUIRED COURSES		31

Elective Courses		Points
<i>Choose 2 or more courses from this list or from alternatives approved by your academic advisor.</i>		
P6503	Introduction to Health Economics	3
P8105	Data Science I	3
P8108 **	Survival Analysis	3
P8109	Statistical Inference	3
P8116	Design of Medical Experiments	3
P8133	Bayesian Analysis and Adaptive Designs in Clinical Trials	3
P8157 **	Analysis of Longitudinal Data	3
P8401	Pharmacoepidemiology	3
G4010	Responsible Conduct of Research and Related Policy Issues	1
GU4200	Biopharmaceutical Development and Regulation	3
W4201	Seminar in Biopharmaceutical Development and Regulation	3

* P8142 may be replaced by P8133, which requires P8104 and P8109

** requires P8104 and P8109



Sample Timeline

Fall I	Spring I	Fall II	Spring II
P6400 Principles of Epidemiology I	P6170 New Drug Development	P6110 Statistical Computing in SAS	P8144 Pharmaceutical Statistics
P8104 Probability	P8120 Analysis of Categorical Data	P8142 Clinical Trial Methodology (or P8133 Adaptive Designs)	P8185 Capstone Consulting Seminar
P8130 Biostatistical Methods	P8140 Introduction to RCTs	P8180 Research Data Coordination	Complete practicum requirements
	Elective	Elective	

Statistical Genetics

Director: Prakash Gorroochurn, PhD

The Master of Science in Biostatistics degree program - Statistical Genetics track (MS/SG) prepares well-qualified students to use advanced modern statistical genetic methods to dissect complicated human genetic archeology with cutting-edge technologies. The high powered curriculum keeps pace with the fast developing field of statistical genetics. Students begin with a rigorous grounding in statistical theory and practice, and then incorporate modern analytic methods into their tool box via new coursework. Students are exposed to the rapidly evolving implementation of those methods in both the hands on practicum experience and consulting work, and in Journal Club reviews of modern methods presented and applied in the literature

Course Requirements

MS/SG students are expected to gain proficiency in genetic study design and analysis as represented by the courses listed below. Students must complete a minimum of 36 academic credits to earn the MS/SG degree, of which 30 points must be taken at the Mailman School of Public Health. *Up to two electives may be taken pass/fail, especially to encourage students to take courses outside their field of experience.*

Note that some courses in the required curriculum may be waived based on prior coursework with approval of the faculty advisor. In this event, the student may substitute another, more advanced course in place of the waived course.

In advance of beginning the MS program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health.

Students' progress will be reviewed after each semester. Those students whose academic performance falls below a cumulative B average (3.0 GPA) in required courses may not be allowed to graduate without remedial course work.

A student is considered full-time in the MS/SG program if he or she takes a minimum of 12 credits per semester.

International students are required to be registered for at least 12 credits during their second and third semesters.

Practicum Requirement

One term of practical experience is required of all students, providing educational opportunities that are different from and supplementary to the more academic aspects of the program. The practicum may be fulfilled during the school year or over the summer. Arrangements are made on an individual basis in consultation with faculty advisors who must approve both the proposed practicum project prior to its initiation, and the report submitted at the conclusion of the practicum experience. Students will be required to make a poster presentation at the department's Annual Practicum Poster Symposium which is held in early May. See the [practicum information section](#) at the back of the handbook for details and links to required forms.

Capstone Experience

A formal culminating experience is required for graduation. The capstone consulting experience is designed to enable students to demonstrate their ability to integrate their academic studies with the role of biostatistical consultant/collaborator, which will comprise the major portion of their future professional practice. Students register for P8185 Capstone Consulting Seminar, a one-semester, one-credit course during their final spring semester. The course requires students to attend at least one session of the Biostatistics Consulting Service and present the consult to the class for discussion. The Consulting Service, which is run by faculty in the Department of Biostatistics, offers advice on data analysis and appropriate methods of data presentation for publications, and provides design recommendations for public health and clinical research, including preparation of grant proposals and manuscripts. Operating five days a week, the Consulting Service is a free biostatistical resource for public health and other health sciences researchers at Columbia University.

Curriculum (TOTAL POINTS: 35 OR MORE)

Required Courses

<i>Core Biostatistics Courses</i>		Points
P6400	Principles of Epidemiology	3
P8104	Probability	3
P8105	Data Science I	3
P8109	Statistical Inference	3
P8130	Biostatistical Methods I	3
P8131	Biostatistical Methods II	3
<i>Core Genetics Courses</i>		
P8119	Adv. Statistical and Computational Methods in Genetics & Genomics	3
P8139	Statistical Genetics Modeling	3
P8149	Human Population Genetics	3
P8185	Capstone Consulting Seminar	1
TOTAL POINTS FROM REQUIRED COURSES		28

Elective Courses

Choose 3 or more courses from this list or from alternatives approved by your academic advisor.

		Points
COMS W1003	Introduction to Computer Science and Programming in C	3
CBMF W4761	Computational Genomics	3
COMS W4771	Machine Learning	3
P8160	Topics in Advanced Statistical Computing	3
P8405	Genetics in Epidemiology	3
P8438	Epidemiology II: Design and Conduct of Observational Epidemiology	3

Sample Timeline

Fall I	Spring I	Fall II	Spring II
P8104 Probability	P8109 Statistical Inference	P6400 Principles of Epidemiology I	P8185 Capstone Consulting Seminar
P8105 Data Science I	P8131 Biostatistical Methods II	P8119 Advanced Statistical & Computational Methods	Complete practicum requirements
P8130 Biostatistical Methods I	P8139 Statistical Genetics Modeling	Elective	
P8149 Human Population Genetics	Elective	Elective	

Theory and Methods

Director: Qixuan Chen, PhD

The Master of Science in Biostatistics degree program - Theory and Methods track (MS/TM) is designed to prepare individuals for a career applying statistical methods in the biomedical sciences. The MS/TM program is the appropriate program for a student whose goal is to work effectively as a biostatistician in a biomedical, clinical, or laboratory research setting; or for a student who plans to pursue a PhD in biostatistics.

Course Requirements

MS/TM students are expected to master certain mathematical and biostatistical concepts and techniques as represented by the courses listed below. Students must complete a minimum of 36 points to earn the MS/TM degree, of which 28 points must be taken at the Mailman School of Public Health. **One selective course and up to two electives may be taken pass/fail (students cannot take more than 9 credits as P/F).**

Note that some courses in the required curriculum may be waived based on prior coursework with approval of faculty advisor. In this event, the student may substitute another, more advanced course in place of the waived course.

In advance of beginning the MS program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health.

Students' progress will be reviewed after each semester. Those students whose academic performance falls below a cumulative B average (3.0 GPA) in required courses may not be allowed to graduate without remedial course work.

A student is considered full-time in the MS/TM program if he or she takes a minimum of 12 credits per semester. **International students are required to be registered for at least 12 credits during their second and third semesters.**

Practicum Requirement

One term of practical experience is required of all students, providing educational opportunities that are different from and supplementary to the more academic aspects of the program. The practicum may be fulfilled during the school year or over the summer. Arrangements are made on an individual basis in consultation with faculty advisors who must approve both the proposed practicum project prior to its initiation, and the report submitted at the conclusion of the practicum experience. Students will be required to make a poster presentation at the department's Annual Practicum Poster Symposium which is held in early May. See the [practicum information section](#) at the back of the handbook for details and links to required forms.

Capstone Experience

A formal culminating experience is required for graduation. The capstone consulting experience is designed to enable students to demonstrate their ability to integrate their academic studies with the role of biostatistical consultant/collaborator, which will comprise the major portion of their future professional practice. Students register for P8185 Capstone Consulting Seminar, a one-semester, one-credit course during their final spring semester. The course requires students to attend at least one session of the Biostatistics Consulting Service and present the consult to the class for discussion. The Consulting Service, which is run by faculty in the Department of Biostatistics, offers advice on data analysis and appropriate methods of data presentation for publications, and provides design recommendations for public health and clinical research, including preparation of grant proposals and manuscripts. Operating five days a week, the Consulting Service is a free biostatistical resource for public health and other health sciences researchers at Columbia University.

Curriculum (TOTAL POINTS: 36 OR MORE)

Required Courses		Points
P6400	Principles of Epidemiology	3
P8104	Probability	3
P8105	Data Science I	3
P8109	Statistical Inference	3
P8130	Biostatistical Methods I	3
P8131	Biostatistical Methods II	3
P8185	Capstone Consulting Seminar	1
TOTAL POINTS FROM REQUIRED COURSES		19

Selective Courses		Points
<i>Choose 1 course from each group.</i>		
GROUP 1		
P8116	Design of Medical Experiments	3
P8140	Randomized Clinical Trials	3
P8142	Clinical Trial Methodology	3
P8144	Pharmaceutical Statistics	3
P8133	Bayesian Analysis and Adaptive Designs in Clinical Trials	3
GROUP 2		
P8108	Survival Analysis	3
P8157	Analysis of Longitudinal Data	3
TOTAL POINTS FROM SELECTIVE COURSES		6

Elective Courses		Points
<i>Choose 4 or more courses from this list or from alternatives approved by your academic advisor.</i>		
P6110	Statistical Computing with SAS	3
P8106	Data Science II	3
P8108	Survival Analysis (if not chosen as a selective)	3
P8119	Adv Statistical and Computational Methods in Genetics & Genomics	3
P8122	Statistical Methods for Causal Inference	3
P8123	Analysis of Health Surveys	3
P8149	Human Population Genetics	3
P8157	Analysis of Longitudinal Data (if not chosen as a selective)	3
P8158	Latent Variable and Structural Equation Modeling for Health Sciences	3
P8160	Topics in Advanced Statistical Computing for Health Sciences	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3
P9120	Topics in Statistical Learning and Data Mining I	3



Sample Timeline

Fall I	Spring I	Fall II	Spring II
P6400 Principles of Epidemiology I	P8109 Statistical Inference	Selective/Elective	P8185 Capstone Consulting Seminar
P8104 Probability	P8131 Biostatistical Methods II	Selective/Elective	Complete practicum requirements
P8105 Data Science	Selective/Elective	Selective/Elective	
P8130 Biostatistical Methods I	Selective/Elective	Selective/Elective	



Patient Oriented Research

Director: Todd Ogden, PhD

The Master of Science in Patient Oriented Research (MS-POR) degree program provides training in the fundamentals of clinical and translational investigation, with a view to enabling young researchers to compete more effectively for research funding. MS/POR students are trained in the design, conduct, and evaluation of clinical research studies, with close supervision and support from the Program Director.

The program is comprised of an interdisciplinary series of courses and colloquia that reflects both the public health faculty's expertise in design and conduct of research studies and the clinical faculty's intimate knowledge of human health and patient care.

Course Requirements

The required courses are intended to enable degree candidates to gain proficiency in study design, application of commonly-used statistical procedures, facility with statistical software packages, and ability to successfully interpret and communicate the results of an analysis. The overall goal is to make graduates more competitive in pursuit of research funding. The two-year MS-POR curriculum consists of 30 credits in total and a culminating Master's Essay. One selective course and up to two electives may be taken pass/fail (students cannot take more than 9 credits as P/F).

MS/POR candidates must begin study during their first summer by enrolling in the **Columbia Summer Research Institute** (CSRI). In the CSRI, students will earn 10 credits, completing courses in biostatistics, epidemiology, NIH grant writing, health disparities research and decision analysis. With this, students earn one-third of the required credits in the first summer, leaving greater flexibility and fewer scheduling commitments over the remaining months.

Note that some courses in the required curriculum may be waived based on prior coursework with approval of a student's faculty advisor. In this event, the student may substitute another, more advanced course in place of the waived course.

In advance of beginning the MS program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health.

Students' progress will be reviewed after each semester. Those students whose academic performance falls below a B average (3.0 GPA) in required courses may not be allowed to graduate without remedial course work.

Capstone Experience

As part of the MS/POR training, each student is required to register for Public Health P9165 Master's Essay – Patient Oriented Research, and complete a master's essay consisting of the construction of an NIH-style grant application. The student is supervised by a Project Sponsor from biostatistics and by a clinical mentor from the student's own field of expertise. At the end of the term, each student will submit a research grant proposal, following NIH guidelines for applications. Each proposal will be reviewed by the program leaders, followed by a formal presentation to the TRANSFORM (Training And Nurturing Scholars For Research that is Multidisciplinary) Advisory Board. The completion, submission, and presentation of the research proposal fulfill the practicum requirement.

Curriculum (TOTAL POINTS: 30 OR MORE)

Required Courses		Points
P6104	Introduction to Biostatistical Methods (CSRI)	3
P6400	Principles of Epidemiology (CSRI)	3
P8103	Colloquium on Patient Oriented Research <i>(taken over four semesters)</i>	2
P8120	Analysis of Categorical Data	3
P8182	Writing a Successful Grant Application (CSRI)	1
P8568	Decision Analysis for Clinical and Public Health Practices (CSRI)	2
P8750	Race and Health (CSRI)	1
P9165	Master's Essay - Patient Oriented Research	0
G4010	Responsible Conduct of Research and Related Policy Issues	1
M9780	Funding for Research Activities: Basic Issues in Obtaining Support	1
89260	Building Interdisciplinary Research Models <i>(also G9260 or N9260)</i>	2

TOTAL POINTS FROM REQUIRED COURSES 19

In addition to the 11 courses listed above, students are required to take at least two of the following courses of which at least one course has to be from the Biostatistics and Epidemiology list.

Restricted electives (selectives)		Points
Precision Medicine and Genetics Electives		
M7208	Precision Medicine	3
P6385	Principles of Genetics and the Environment	3
P8119	Adv Statistical & Computational Methods in Genetics & Genomics	3
P8405	Genetics in Epidemiology	3
Mechanisms/Molecular Electives		
G4500	Cancer Biology 1	3
G6003	Mechanisms in Human Disease I	4.5
P8307	Molecular Epidemiology	3
P8308	Molecular Toxicology	3
P8312	Fundamentals of Toxicology	3
P8319	Biological Markers of Chemical Exposure	3
Data and Computing Electives		
G4001	Introduction to Computer Application in Health Care & Biomedicine	3
P6110	Statistical Computing Using SAS	3

*MSPOR elective courses continued
on next page*

curriculum CONTINUED

P8101	Introduction to Health Data Science	3
P8105	Data Science I	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3
P8451	Introduction to Machine Learning for Epidemiology	3
	Dissemination & Implementation Science and Community-based Participatory Research Electives	
P8792	Dissemination and Implementation Science	3
P8771	Community Based Participatory Research	3
	Biostatistics and Epidemiology Electives	
P8100	Applied Regression I	3
P8110	Applied Regression II	3
P8112	Systematic Review and Meta-analysis	1.5
P8122	Statistical Methods for Causal Inference	3
P8140	Randomized Clinical Trials	3
P8142	Clinical Trials Methodology	3
P8400	Epi III: Applied Epidemiological Analysis	3
P8401	Pharmacoepidemiology	3
P8438	Epi II: Design and Conduct of Observational Epidemiology	3
P8450	Clinical Epidemiology	3
P8777	Survey Research Methods	3
P8902	Introduction to Mixed Methods	3

Elective Courses. Students will choose elective courses from one or more of the following:

- Departments of Epidemiology or Biostatistics, or other departments at the School of Public Health
- From the list of restricted electives (selectives) – see above
- Elective courses from other Columbia schools – see list below

Electives in Clinical and Translational Research outside of Mailman School of Public Health

<i>Electives in Biostatistics</i>		Points
B8128	Healthcare Investment and Entrepreneurship	1.5
B8342	Healthcare Investment and Deal-making	1.5
B8692	Pharmaceutical Drug Commercialization: Strategy & Practice	1.5
B8745	Forecasting for Drug Development Strategy	1.5
E6893	Topics in Information Processing: Big Data Analytics	3
G4006	Translational Bioinformatics	3
G4062	Public Health Informatics	1



Sample Timeline

Summer I	Fall I	Spring I	Fall II	Spring II
P6104 Intro to Biostatistical Methods	P8103 Colloquium (0.5)	P8103 Colloquium (0.5)	P8103 Colloquium (0.5)	P8103 Colloquium (0.5)
P6400 Principles of Epidemiology I	Elective	P8120 Analysis of Categorical Data	Elective	M9780 Funding for Research Activities
P8182 Writing a Successful NIH Grant	Selective	G4010 Responsible Conduct of Research	P9165 Master's Essay (POR Capstone)	
P8568 Decision Analysis		Selectives/Elective		89260 Building Interdisciplinary Research Models
P8750 Race and Health				

Doctoral degree programs

The Department of Biostatistics offers two doctoral degree programs: the Doctor of Public Health (DrPH) and the Doctor of Philosophy (PhD). Both the DrPH and PhD programs train candidates to apply state-of-the-art statistical methods to the analysis of important public health issues and potential solutions, but differ in their relative emphasis on application versus statistical theory.

Doctor of Public Health (DrPH)

DrPH training places relatively greater emphasis on the application of statistical methods to public health problems, although many DrPH students propose new methods and contribute to the advancement of statistical theory as part of their dissertation research.

Upon satisfactory completion of the DrPH degree in Biostatistics, graduates will be able to:

Data Analysis and Computing

- Identify and implement advanced statistical models for the purposes of estimation, comparison, prediction, and adjustment in non-standard settings.

Public Health and Collaborative Research

- Describe the foundations of public health, including the biological, environmental, behavioral, and policy factors that affect the health of populations;
- Develop and execute calculations for power and sample size when planning research studies with complex sampling schemes; and
- Formulate and prepare a written statistical plan for analysis of biomedical or public health research data that clearly reflects the research hypotheses of the proposal in a manner appropriate for scientists with varied backgrounds.

Consulting

- Function as an effective consultant in biomedical and public health research projects; and
- Communicate and write effectively in order to describe complex topics in a consulting environment.

Data Management

- Identify the uses to which data management can be put in practical statistical analysis, including the establishment of standards for documentation, archiving, auditing, and confidentiality; guidelines for accessibility; security; structural issues; and data cleaning;
- Differentiate between analytical and data management functions through knowledge of the role and functions of databases, different types of data storage, and the advantages and limitations of rigorous database systems in conjunction with statistical tools; and
- Assess database tools and the database functions of statistical software, with a view to explaining the impact of data management processes and procedures on their own research.

Teaching

- Explain and illustrate principles of study design and data analytic techniques to public health students enrolled in first and second level graduate public health courses.

Biostatistical Research

- Identify and integrate new developments in the statistical literature for challenging research problems in public health; and
- Generate efficient computer code to implement sophisticated statistical techniques.



Doctor of Philosophy (PhD)

PhD training places relatively greater emphasis on the development of novel statistical theory and methods. A PhD dissertation must represent an original contribution to statistical theory or methods that has relevance to a real biomedical or public health application.

Upon satisfactory completion of the PhD degree in Biostatistics, graduates will be able to:

Data Analysis and Computing

- Identify and implement advanced statistical models for the purposes of estimation, comparison, prediction, and adjustment in non-standard settings.

Public Health and Collaborative Research

- Develop and execute calculations for power and sample size when planning research studies with complex sampling schemes;
- Formulate and prepare a written statistical plan for analysis of biomedical or public health research data that clearly reflects the research hypotheses of the proposal in a manner appropriate for scientists with varied backgrounds; and
- Evaluate research reports and proposals for research funding on the basis of their scientific integrity, validity, and the strength of the quantitative analysis.

Consulting

- Function as an effective consultant in biomedical and public health research projects; and
- Develop communication and writing skills in a consulting environment.

Data Management

- Identify the uses to which data management can be put in practical statistical analysis, including the establishment of standards for documentation, archiving, auditing, and confidentiality; guidelines for accessibility; security; structural issues; and data cleaning;
- Differentiate between analytical and data management functions through knowledge of the role and functions of databases, different types of data storage, and the advantages and limitations of rigorous database systems in conjunction with statistical tools; and
- Describe the different types of database management systems, the ways these systems can provide data for analysis and interact with statistical software, and methods for evaluating technologies pertinent to both.

Teaching

- Explain and illustrate selected principles of study design, probability theory, inference, and data analytic techniques to public health students enrolled in first and second level graduate public health courses; and
- Explain advanced concepts in the theory of statistical inference to graduate students in biostatistics and mathematical statistics.

Biostatistical Research

- Identify and integrate new developments in the statistical literature for challenging research problems in biomedicine and public health; and
- Generate efficient computer code to implement sophisticated statistical techniques; and
- Recognize gaps in current inferential methods that limit further public health research and propose and develop solutions based on rigorous theoretical considerations.



Doctor of Public Health

Director: Shing Lee, PhD

The Doctor of Public Health degree in Biostatistics (DrPH) prepares candidates to apply modern statistical methods to the solution of important public health problems as leaders of multidisciplinary research teams. The degree program is administered by the Standing Doctoral Committee of the Mailman School of Public Health, which carries out faculty policy on admission to the doctoral program and upholds the criteria for granting the degree.

Course Requirements

The Doctor of Public Health degree calls for completion of an approved program of study totaling no less than 36 doctoral credits. Upon completion of 36 credits of coursework, a student is permitted to take the written qualifying examination. In some instances it may be determined by the Department that a student needs more than 36 post-MPH course credits before the qualifying examination.

DrPH students must maintain continuous registration every semester from the start of the program until deposit of the doctoral dissertation. After completion of all coursework students register for Doctoral Registration (RSRHP0001) each term until they are ready to graduate.

No more than 10 credits may be tutorials, and no more than six may be earned at 6000-level courses at the Mailman School of Public Health or 4000-level courses at the Graduate School of Arts and Sciences; the Department may apply to the Doctoral Committee for a variance on the six-credit rule on a case-by-case basis.

A list of required courses is given below. A DrPH student who has not previously earned an MPH degree must complete courses in each of the core areas of public health: biostatistics, environmental health sciences, epidemiology, health policy, and social and behavioral sciences. Core courses will have to be completed through the Columbia MPH Core or completed by taking a comparable graduate level course at an accredited institution. These comparable courses must be approved by the School's Doctoral Committee. A DrPH student who has not previously earned an MPH will also be required to take the online course, PUBH P6025-Introduction to Public Health. The credits accrued for completing the core requirements DO NOT count towards the 36 doctoral credits.

A grade of B or better is necessary in all required courses. Up to 2 elective courses may be taken pass/fail.

Training in Interdisciplinary Research

The curriculum is designed to enable students to integrate their training in statistical methods and theory with the role of biostatistical consultant/collaborator on interdisciplinary teams, which will comprise a major portion of their future professional practice. Statistical Practices and Research for Interdisciplinary Sciences I & II are courses in which students gain experience with design, data analysis, and both oral and written presentation communication through exposure to several consulting projects. PhD students are required to enroll in P9185 prior to taking the Applications section of the Qualifying Exam and required to enroll in P9186 after taking the Applications Qualifying Exam. Students with extensive work experience in the field may request to waive the P9186 requirement from the Director of the DrPH Program.

Curriculum (TOTAL POINTS: 36 OR MORE)

Required Courses		Points
P8104*	Probability	3
P8105*	Data Science I	3
P8106*	Data Science II	3
P8108*	Survival Analysis	3
P8109*	Statistical Inference	3
P8120*	Analysis of Categorical Data	3
P8130*	Biostatistical Methods I	3
P8131*	Biostatistical Methods II	3
P8157*	Analysis of Longitudinal Data	3
P9185	Statistical Practices and Research for Interdisciplinary Sciences I	3
P9186	Statistical Practices and Research for Interdisciplinary Sciences II	1.5
P9070	DrPH Case Studies in Public Health Leadership I & II	3
P9050	DrPH Seminar in Strategic Management	1.5
P9040	DrPH Seminar in Management and Organizational Behavior	1.5
P9060	Essentials of Teaching and Communication for Doctoral	1.5
Elective Courses		Points
P8142	Clinical Trial Methodology	3
P8116	Design of Medical Experiments	3
P8122	Statistical Methods for Causal Inference	3
P8123	Analysis of Health Surveys	3
P8140	Introduction to Randomized Clinical Trials	3
P8158	Latent Variable and Structural Equation Modeling for Health Sciences	3
P8160	Topics in Advanced Statistical Computing	3
P8180	Relational Databases and SQL Programming for Research and Data Science	3
P9120	Topics in Statistical Learning and Data Mining I	3
89260	Building Interdisciplinary Research Models	2

**requirements can be waived, consult with your faculty advisor*

Departmental Colloquium

All doctoral students are **required** to attend the Departmental Colloquium and Research Talks held weekly each semester. Dates, times, and locations will be posted on the Department electronic board, as well as on the Department's website and Facebook page.

The Applied Practice Experience (Practicum) for DrPH Students

Regardless of the amount or level of prior experience, all DrPH students are required to engage in an applied practice experience in which students are responsible for completion of at least one project that is meaningful for an organization and to advance public health practice.

The work product may be a single project or a set of related projects that demonstrate a depth of competence. The deliverable must contain a reflective component that includes the student's expression of personal and/or professional reactions to the applied practice experience. This may take the form of a manuscript, journal article or other written product, a professional portfolio, or another deliverable that serves to assess the ability of the student to meet department and School competencies.

The applied practice experience should take place within an organization external to the student's school or program so that it is not merely an academic exercise, but application of learning to a "real world" setting. Relevant organizations may include governmental, non-governmental, non-profit, industry and for-profit settings. The Office of Field Practice and individual departments identify sites in a manner that is sensitive to the needs of the agencies or organizations involved, and sites should benefit from Mailman students' experiences. The applied practice experience may be completed within a student's own work setting, as long as the applied practice experience differs substantially from a student's current job description and meets the required competencies described below.

The applied practice experience must meet a minimum of five (5) foundational and/or concentration-specific competencies that are reinforced and/or assessed through application. One of these competencies must be a school-wide or a departmental-specific competency in leadership, management, and governance. Competencies for the applied practical experience must be agreed upon by the student, advisor, and applied learning experience preceptor, as specified in the statement of work form.

While there is not a minimum number of hours for the applied practice experience, it does require substantive, quality opportunities that address the identified competencies.

Qualifying Examination

There is a two-part qualifying examination for all DrPH candidates in Biostatistics that must be completed before going on to the oral comprehensive examination.

Basic Inference. The first part assesses basic familiarity with statistical inference as presented in the course P8109 Statistical Inference. Students who have taken this course (or a comparable graduate course) and have received a grade of B+ or above automatically satisfy the basic inference requirement. All others will be required to take a written examination testing their knowledge of the material in this course. In all cases, students must fulfill the requirement within two years of starting the doctoral program. Students must pass the basic inference requirement before they may sit for the Applications exam.

Applications. The Applications portion covers the practical analysis of data. This part focuses on addressing applied problems requiring statistical inference based on data analysis. The purpose of the Applications exam is to ensure that the student is able to determine the appropriate statistical and analytic approaches needed to solve real world public health / medical problems, correctly interpret the statistical results from these approaches, and translate and summarize those findings into language that public health and medical professionals would find useful.

Course Work and Progressing toward the Applications portion. Preparation should include additional coursework in skills classes, review and thorough understanding of the material in the suggested readings, group and individual study sessions, completion of timed practice tests, and enrollment in P9185 Statistical Practices and Research for Interdisciplinary Sciences I, a course in which students gain exposure to real world design, analysis, and report writing.

With approval and consent of his or her academic advisor, the student should inform the Director of Academic Programs two months in advance of sitting for the Applications portion of the exam.

Grading on the Applications exam is holistic, and can also take into account performance in coursework, and other factors deemed relevant. A score below 65% on the exam would generally be considered unsatisfactory. The student will be allowed no more than two attempts at passing either the Basic Inference or Applications parts of the examination. The Applications portion must be taken and passed by the end of the third year in the DrPH program.

Questions from prior years are available to the student to assist in preparing for the examination.

Reading List

The following list consists of textbooks that are used in the courses required for the DrPH degree, plus additional references which are generally at the appropriate level for the DrPH Qualifying Examinations. Those marked with an asterisk are highly recommended to students preparing for their examinations.

- Breslow NE and Day NE, *Statistical Methods in Cancer Research*
Conover WJ, *Practical Nonparametric Statistics*
Cox DR and Oakes D, *Analysis of Survival Data*
Fleiss JL, *The Design and Analysis of Clinical Experiments*
* Fleiss JL, Levin B, and Paik MC, *Statistical Methods for Rates and Proportions*
Hogg RV and Craig AT, *Introduction to Mathematical Statistics*
* Hosmer D and Lemeshow S, *Applied Logistic Regression*
* Johnson RA and Wichern DW, *Applied Multivariate Statistical Analysis*
Kalbfleisch JD and Prentice RL, *Statistical Analysis of Failure Time Data*
Kleinbaum DG and Kupper LL, *Applied Regression Analysis and other Multivariable Methods*
* Lawless JF, *Statistical Models and Methods for Lifetime Data*
* Lee ET, *Statistical Methods for Survival Data Analysis*
Lehmann ER, *Nonparametrics: Statistical Methods Based on Ranks*
Mardia KV, Kent JT, and Bibby JM, *Multivariate Analysis*
* Mood AM, Graybill FA, and Boes D, *Introduction to Statistical Inference*
Morrison DF, *Multivariate Statistical Methods*
* Mosteller F and Tukey JW, *Data Analysis and Regression*
* Neter J, Wasserman W, and Kutner MH, *Applied Linear Statistical Models*
Rao CR, *Linear Statistical Inference and Its Applications*
Scheffe H, *The Analysis of Variance*
Searle SR, *Linear Models*
Snedecor GW and Cochran WG, *Statistical Methods*
Tukey JW, *Exploratory Data Analysis*

Oral Comprehensive Examination

After completing all course work and passing the qualifying examination described above, the DrPH candidate begins planning for the Integrative Learning Experience (ILE). The oral comprehensive examination for the DrPH in Biostatistics is intended to examine the student's mastery of the current state of knowledge about his or her project area, and thus to indicate whether the student is prepared to undertake such a project. **The Oral Comprehensive Examination should be taken no later than six months after passing the qualifying exam.**

Composition of the Examining Committee. The examining committee will consist of five members approved by the chair of the Doctoral Program Subcommittee on Biostatistics, and will include:

- i) three members who are inside examiners (i.e. holding a formal appointment or approved as a dissertation sponsor);
- ii) preferably two (but at least one) members who are outside examiners.

The latter faculty should represent disciplines closely related to the area of application of the student's proposed research. **After the sponsor obtains consent from each member, the *faculty sponsor* submits the list of names to the Chair of the Department and to the Chair of the Departmental Subcommittee on Biostatistics (DPSOB) for approval**, who then recommends the student's committee to the DrPH Committee of the Mailman School of Public Health.

Scheduling the Exam. The oral comprehensive examination should be taken within one year of passing the qualifying examination.

Nature of the Examination. After the committee selection and approval process has been completed, the student submits in writing a description of the current state of knowledge about the proposed area of research. This submission should be from 15 to 25 pages in length and contain between 15 and 20 references. This paper serves as the basis for the oral comprehensive examination. The student must give each member of the Examining Committee this written submission and discuss with each any additions or deletions that the committee member feels should be incorporated in the write-up. Since the final written submission and the references therein will constitute the basic material upon which the student will be examined, each member of the committee and the student must come to an agreement on the scope of the submission. It should be neither too narrow nor too broad in scope. After all members of the ad hoc committee approve the submission, the examination is scheduled within the next 60 days. The written submission may contain original results by the student, but this is not required.

ORAL COMPREHENSIVE EXAMINATION CONTINUED

Format of the Exam. The actual examination shall be an Oral Comprehensive Examination conducted by the Examining Committee as follows:

1. The chair of the Examining Committee will not be the ILE advisor but another member of the ad hoc committee.
2. The examination will run approximately two hours and will consist of an oral presentation of the content of the written submission by the student (a planned presentation of about 30 minutes is appropriate), which may be interrupted by members of the Examining Committee with appropriate questions on the material presented or relevant related material. The chair of the Examining Committee may challenge any question felt to be unrelated to the written submission and its background material.
3. After the presentation and questions, each member may ask additional questions of the examinee. Any such questions should be within the broad content of the written submission and its references. Again, the Examining Committee chair may challenge any question felt to be too far removed from the basic material upon which the examination is based, namely on the written submission and the references therein.
4. After all questions are completed, the examinee leaves the room and the committee then votes on whether or not the examinee passed the examination. All members must agree in order for the student to pass the examination. Instead of pass or fail, the committee may unanimously decide upon the option of retesting the student within a six-month period on the same written submission.

The committee's decision will be put into writing by the chair of the Examining Committee, as well as brief comments on the strengths and weaknesses of the student's performance as deemed necessary. Copies of this statement will be sent to the student and placed in the student's file.

Second Attempt at Passing. The student is entitled to no more than two attempts at passing the Oral Comprehensive Examination. The second attempt need not be based on the same written submission nor be examined by the same committee, but the same rules will govern the second attempt, including approval by the committee of the written submission. The second attempt must be made no more than 6 months after the first attempt.

Upon passing the Oral Comprehensive Examination, a student will typically ask his sponsor or another member of the faculty to agree to serve as the student's sponsor. No formal approval of an ILE topic is required; however, a suitable and mutually agreeable topic must be established by the student and advisor. As stated earlier, it is often the case that the Oral Comprehensive Examination is on a topic related to the student's ILE, although this is not a formal requirement.

Progressing toward the Integrated Learning Experience Defense

Between the Oral Exam and the ILE Defense, the DrPH student is required to present his/her project in two public settings. The first is the Doctoral Research Seminar, usually held in the spring, where doctoral students present their work to the faculty and their peers. The second setting is the preparation and presentation of a paper (or poster) at a conference of professional societies or at a statistics or biostatistics departmental talk for job interviews. A select, but not exhaustive, list of such societies is presented below. More information is available on the Doctoral Bulletin Board. Travel funds are often available.

Example of Professional Societies / Associations:

- *American Statistical Association (ASA)*
- *American Public Health Association (APHA)*
- *International Biometric Society (ENAR/IBS)*
- *Joint Statistical Meetings (JSM)*
- *Society for Clinical Trials (SCT)*

The Integrated Learning Experience (ILE)

Once a DrPH student has advanced to doctoral candidacy, s/he begins to develop a proposal for the ILE project. The topic must deal with an important problem or issue in public health which can be addressed by the sound and original application of existing statistical methods. It must demonstrate that the candidate has engaged in independent and original research that has advanced our understanding of or knowledge about the public health problem, though the methods themselves need not be original. After the project is successfully defended, the doctoral degree is awarded by the Mailman School of Public Health in the Faculty of Medicine.

In most cases, completion of DrPH course work and written qualifiers should take no more than two full-time academic years. On average, the ILE may take an additional two or three full-time academic years. An overall time limit of seven years is set from the date of first registration as a doctoral student.

In unusual instances a student may wish to change ILE sponsors, for instance, if the student's project requires different areas of expertise than originally anticipated. In such cases the student may seek approval from a new faculty sponsor. The candidate must inform the Department Chair and the previous sponsor that the new sponsor will assume the previous sponsor's duties. At this point the student may also decide to pursue a new project topic, with approval of the new sponsor, but in all cases the rules governing time limits and extensions remain in force.

DrPH candidates are required to submit an electronic copy of their final report to the department. Hard copies of past reports are available in the department administrative office (Room 626).

For more details regarding the DrPH ILE, the student is referred to the Mailman Student Handbook: www.mailman.columbia.edu/people/current-students/academics/student-handbooks.

Some Past DrPH ILE Titles

The titles below are provided to give students some idea of ILE topics which in past years have proved appropriate for the DrPH degree:

Clustering Algorithm for Zero-Inflated Data, Anusorn Thanataveerat (2020)

Statistical Issues in Platform Trials with a Shared Control Group, Jessica Overbey (2019)

Bayesian Modeling of Latent Heterogeneity in Complex Survey Data and Electronic Health Records, Rebecca Anthopolos (2019)

Statistical Methods for the Study of Etiologic Heterogeneity, Emily Zabor (2019)

Statistical Methods for Integrated Cancer Genomic Data Using a Joint Latent Variable Model, Esther Drill (2018)

Bayesian Modeling for Mental Health Surveys, Sharifa Williams (2018)

Data-Driven Methods for Identifying and Validating Shorter Symptom Criteria Sets: The Case for DMS-5 Substance Use Disorders, Cheryl Raffo (2018)

Design and Analysis of Sequential Multiple Assignment Randomized Trial for Comparing Multiple Adaptive Interventions, Xiaobo Zhong (2018)

Prognostic Modeling in the Presence of Competing Risks: An Application to Cardiovascular and Cancer Mortality in Breast Cancer Survivors, Nicole Leoce (2016)

New Estimating Equation Approach for the Secondary Trait Analyses in Genetic Case-Control Studies, Xiaoyu Song (2015)

Identifying Patterns in Behavioral Public Health Data Using Mixture Modeling with an Informative Number of Repeated Measures, Gary Yu (2013)

A Life Expectancy-based Comprehensive Quantification of Structural-level Health Disparities, Emma Benn (2012)

An Index of Aging-Relatedness with Relevance to Genetic and Environmental Contributions to Mortality and Disease Incidence in a Population, Gilberto Levy (2011)

The Familial Aggregation of Epilepsy, Anna Peljto (2010)

Analysis of the MT CT-Plus Initiative: An Application of a Piecewise Multilevel Latent Variable Regression Model, Shean-Sheng Wang (2010)

Analyses on Double-Blinded Fetal Tissue Transplant Study on Patients with Severe Parkinson's Disease and Reliability Assessment on Outcome Measurements, Richard Kao (2010)

Stepwise Procedures for Dose Finding in an Adaptive Clinical Trial of Early Rehabilitation After Acute Stroke, Xi Wu (2010)

Instrumental Variable Estimation for Survival Data: Evaluating the Effectiveness of Radiation Therapy for the Treatment of Lung Cancer in the Elderly in the Presence of Allocation Bias, Juan P. Wisnivesky (2009)



Doctor of Philosophy

Director: Jeff Goldsmith, PhD

The Doctor of Philosophy in Biostatistics (PhD) prepares candidates for leadership roles in the development and application of statistical methods to biomedical research for the advancement of public health. The PhD is awarded by the Graduate School of Arts and Sciences (GSAS) as governed by the Doctoral Program Subcommittee on Biostatistics. The program is administered by the faculty and staff of the Mailman School of Public Health.

Course Requirements

Students take courses in the department of biostatistics, and other academic units representing various fields of application and/or related background material. A student should plan his or her course work in consultation with his or her academic advisor and/or the PhD subcommittee chair. Students wishing to waive one or more required courses must request approval in writing from their faculty advisor and the Director of Academic Programs.

A grade of B or better is necessary in all required courses, except for P9111 which requires a B+ or better. Electives may be taken pass/fail, in order to encourage candidates to take courses outside his or her field of experience.

In advance of beginning the PhD program, any student who has not completed an MPH will be required to take the online course required by the Mailman School: PUBH P6025-Introduction to Public Health.

Training in Interdisciplinary Research

The curriculum is designed to enable students to integrate training in statistical methods and theory with the role of biostatistical collaborator on interdisciplinary teams, which will comprise a major portion of their future professional practice. Statistical Practices and Research for Interdisciplinary Sciences I & II are courses in which students gain experience with design, data analysis, and both oral and written communication through exposure to several consulting projects. PhD students are required to enroll in P9185 during the spring semester prior to taking the take home component of the Qualifying Exam. Students are required to enroll in P9186 during the fall semester after taking the take home component of the Qualifying Exam.

Statistical Inference Problem Seminar

To prepare for the written component of the Qualifying Exam, students are required to take the problem seminar in which students work on problems and discuss problem solving strategy useful for theoretical questions. The problem seminar is held months prior to the written/theory and method qualifying exam.

GSAS Requirements

In addition to registering for individual courses, PhD students are required to register for the Residence Unit (RU) which provides the basis for tuition charges and provides full-time status. Six RUs are required for the PhD degree. RUs may only be earned during fall and spring semesters, not during the summer. PhD students must register for 1 RU each semester up to the total required 6 RUs. After one year of study, students who enter with a Master's degree may apply for advanced standing of two residence units representing work completed in their Master's program. After the student has satisfied the residency requirement they must register for full-time Matriculation & Facilities status until a successful dissertation defense.

Curriculum

Required Courses		Points
P6400*	Principles of Epidemiology	3
P8104*	Probability	3
P8105*	Data Science I	3
P8106*	Data Science II	3
P8109*	Statistical Inference	3
P8130*	Biostatistical Methods I	3
P8131*	Biostatistical Methods II	3
P8160*	Topics in Advanced Statistical Computing	3
P9104	Probability for Biostatisticians	3
P9109	Theory of Statistical Inference I	3
P9110	Theory of Statistical Inference II	3
P9111	Asymptotic Statistics	3
P9120	Topics in Statistical Learning and Data Mining I	3
P9130	Advanced Biostatistical Methods I	3
P9185	Statistical Practices and Research for Interdisciplinary Sciences I	3
P9186	Statistical Practices and Research for Interdisciplinary Sciences II	1.5

Sample Elective Courses		Points
P8108	Survival Analysis	3
P8116	Design of Medical Experiments	3
P8122	Statistical Methods for Causal Inference	3
P8123	Analysis of Health Surveys	3
P8133	Bayesian Analysis and Adaptive Designs in Clinical Trials	3
P8140	Introduction to Randomized Clinical Trials	3
P8142	Clinical Trial Methodology	3
P8144	Pharmaceutical Statistics	3
P8157	Analysis of Longitudinal Data	3

**requirements can be waived if taken a comparable course at the Master's level; consult with your faculty advisor*

Departmental Colloquium

All doctoral students are **required** to attend the Departmental Colloquium and Research Talks held weekly each semester. Dates, times, and locations will be posted on the Department electronic board, as well as on the Department's website and Facebook page.

Qualifying Examination

There is a two-part qualifying examination for all PhD candidates in Biostatistics that must be completed prior to the oral comprehensive examination. The written and take home portions of the exam are to be taken during the same summer semester.

Written Portion - Theory and Methods. The written or theory and method exam draws from material presented in the following courses: P9104, P9109, P9110, P9130, P8104, P8109, P8130 and P8131. The purpose of the written exam is to ensure that the PhD student is able to fully understand and use the mathematical and theoretical tools that form the basis of doctoral level biostatistical research. The exam requires solutions to five questions. Students entering with a Bachelor's are expected to take the exam after their second year; students entering with a relevant Master's are expected to take the exam the summer after their first year.

Course Work and Progressing toward the written portion of the Qualifying Exam. Preparation should include coursework or mastery of content of the material in the required courses, review and thorough understanding of the material in the suggested readings, group and individual study sessions, and completion of timed practice tests. With approval and consent of the student's academic advisor, the student should inform the Director of Academic Programs two months in advance of sitting for the written portion of the Qualifying Exam.

Take Home Portion - Applications. The take home exam covers the practical analysis of data. The examination focuses on applied problems requiring statistical inference based on data analysis, with particular emphasis on material from P8105, P8106, P8130, P8131, P9130 and P9185.

The purpose of the take-home exam is to ensure that the student is able to determine the appropriate statistical and analytic approaches needed to solve real world public health / medical problems, correctly interpret the statistical results from these approaches, and translate and summarize those findings into language that public health and medical professionals would find useful. The take-home exam is administered over a two-day period. Students are encouraged to use personal laptops and any familiar software. Students entering with a bachelors are expected to take the exam after their second year; students entering with a relevant Master's are expected to take the exam the summer after their first year.

Course Work and Progressing toward the take home portion of the Qualifying Exam. Preparation should include additional coursework in skills classes, review and thorough understanding of the material in the suggested readings, group and individual study sessions, completion of timed practice tests, as well as enrollment in P9185. With approval and consent of his or her academic advisor, the student should inform the Director of Academic Programs two months in advance of sitting for the take-home exam.

Grading on the Qualifying Exam is holistic, taking into account performance in coursework, on both portions, and other factors deemed relevant. A score below 65% on either the written or take home portion will generally be considered unsatisfactory. The student will be allowed no more than two attempts at passing either part of the exam. It is strongly recommended that the second attempt be made at the time of the next exam offering.

Exam questions from prior years are available to the student to assist in preparing for the examination.

Reading List

The following list consists of textbooks that are generally appropriate to use for preparing for the PhD qualifying examination.

- Agresti A, *Categorical Data Analysis*
- Bickel PJ and Doksum KA, *Mathematical Statistics*
- Casella G and Berger RL, *Statistical Inference*
- Cox D and Hinkley DV, *Theoretical Statistics*
- Efron B and Tibshirani R, *An Introduction to the Bootstrap*
- Hastie T, Tibshirani R, and Friedman J, *The Elements of Statistical Learning*
- Hettmansperger TP and McKean JW, *Robust Nonparametric Methods*
- Hollander M, *Nonparametric Statistical Methods*
- Johnson RA and Wichern DW, *Applied Multivariate Statistical Analysis*
- Klein JP and Moeschberger ML, *Survival Analysis*
- Lehmann EL, *Point Estimation*
- Lehmann EL, *Testing Statistical Hypotheses* Lehmann EL, *Elements of Large-Sample Theory*
- McCullagh P and Nelder JA, *Generalized Linear Models*
- Rao CR, *Linear Statistical Inference and Its Applications*
- Robert CP and Casella, G, *Monte Carlo Statistical Methods*
- Ruppert D, Wand MP, and Carroll R, *Semiparametric Regression*
- Shao J, *Mathematical Statistics*
- Wickham, H and Golemund, G, *R for Data Science*

Oral Comprehensive Examination

After completing all course work and passing the two-part qualifying examination described in the previous sections, the PhD candidate begins planning for dissertation research and preparing for the Oral Exam. The Oral Comprehensive Examination is intended to demonstrate the student's mastery of the material in a defined statistical content area by verbally presenting a thorough description of the state of the art in that area, identifying limitations or areas of incomplete knowledge in that area, and proposing the development of new methods that would advance that area. This topic area may or may not end up being the student's dissertation topic. **The Oral Comprehensive Examination should be taken no later than six months after passing the two-part qualifying exams.** *Fellows in the program, please note all tuition expenses incurred as a result of any delay in scheduling this exam shall be the responsibility of the student and not the Department of Biostatistics.*

Composition of the Examining Committee. The examining committee will consist of five members approved by the chair of the Doctoral Program Subcommittee on Biostatistics, and will include:

- i) three members who are inside examiners (i.e. holding a formal appointment or approved as a dissertation sponsor);
- ii) preferably two (but at least one) members who are outside examiners.

The chair of the Examining Committee will be a member of the Doctoral Program Subcommittee On Biostatistics. One member of this committee should be the faculty member who acts as the student's sponsor and anticipated thesis advisor. **With the consent of the members of the proposed committee, the faculty sponsor then submits their names for approval by the Chair of the Doctoral Program Subcommittee on Biostatistics.**

Nature of the Examination. After the committee selection and approval process has been completed, the student submits in writing a description of the current state of knowledge about the proposed area of research. This submission should be from 15 to 25 pages in length and contain between 15 and 20 references. This paper serves as the basis for the oral comprehensive examination. The student must give each member of the committee this written submission and discuss with each any additions or deletions that the committee member feels should be incorporated in the write-up. Since the final written submission and the references therein will constitute the basic material upon which the student will be examined, each member of the committee and the student must come to an agreement on the scope of the submission and references. After such modifications to the written submission have been approved by all four members of the Examining Committee, the Comprehensive Exam is scheduled within the next 30 days. The written submission may contain original research by the student, but need not be original in content. It should not be too narrow in scope and should reflect the necessary basic material relevant to the student's chosen area of research. Before and during the examination, the three faculty examiners other than the student's sponsor make suggestions for and may insist on changes in the student's perception of the topic. Part of the student's written submission is an enumeration of as yet unanswered questions. The examiners make their opinions plain as to how important and challenging they perceive these questions to be.

ORAL COMPREHENSIVE EXAMINATION CONTINUED

Format of the Exam. The actual examination shall be an Oral Comprehensive Examination conducted by the Examining Committee as follows:

1. The chair of the Examining Committee will not be the dissertation advisor but some other member of the ad hoc committee.
2. The examination will run approximately two hours and will consist of an oral presentation of the content of the written submission by the student (a planned presentation of about 30 minutes is appropriate), which may be interrupted by members of the Examining Committee with appropriate questions on the material presented or related material. The chair of the Examining Committee may overrule any question felt to be unfair or unrelated to the written submission and its background material.
3. After the presentation and questions, each member may ask additional questions of the examinee. Such questions should be within the broad scope of the written submission and references. Again, the Examining Committee chair may rule against any questions felt to be too far removed from the basic material upon which the examination is based, that is, the written submission and the references therein.
4. After all questions are completed, the examinee leaves the room and the committee then votes on whether or not the examinee passed the exam. Three of the four members must vote to pass the student in order for the student to pass the exam.

The committee's decision will be put into writing by the chair of the Examining Committee, as well as brief comments on the strengths and weaknesses of the student's performance as deemed necessary. Copies of this statement will be sent to the student and placed in the student's file.

Second Attempt at Passing. The student is entitled to no more than two attempts at passing the Oral Comprehensive Examination. The second attempt need not be based on the same written submission nor be examined by the same committee, but the same rules will govern the second attempt, including approval by the committee of the written submission. The second attempt must be made no more than 6 months after the first attempt.

The examination and written submission are designed to focus the examination on basic material which is important to the student's area of research, and allow the Examining Committee to judge that the student fully comprehends this material. Upon passing the Comprehensive Examination, a student will typically ask his sponsor or another member of the PhD subcommittee to serve as the student's dissertation advisor and sponsor. No formal approval of a dissertation topic is required; however, a suitable and mutually agreeable topic must be established by the student and advisor. While it is usually the case that the Oral Comprehensive Examination is on a topic that will become the student's dissertation topic, this is not a formal requirement.

Advancement of PhD Students to the Master of Philosophy Degree

Upon the student's passing the qualifying and oral comprehensive examinations and the successful completion of six residence units beyond the Master's degree (two residence units awarded for a completed Master's degree), he or she is awarded the Master of Philosophy degree. Failure on the Oral Comprehensive Examination implies that it is the Subcommittee's judgment the student is not yet prepared to carry out original research. The awarding of the Master of Philosophy to a student, on the other hand, certifies that the student has mastered the fundamental material necessary for him or her to conduct research in biostatistics. Students who apply for and receive two residence units of advanced standing are still required to complete four additional residence units before the Master of Philosophy may be awarded.

Progressing toward the Dissertation Defense

Between the Oral Exam and the Dissertation Defense, the PhD student is required to present his or her research in a two public settings. Typically, one of these settings is at a Graduate Research Seminar, where doctoral students at various stages of their research present their work to their peers. A second setting is the preparation and presentation of a paper (or poster) at a conference of professional societies.

A select, but not exhaustive, list of such societies is presented below.. Students who are selected to present at a conference can apply for travel funds at the School and department levels. Information requests about available travel funds should be directed to the Director of Academic Programs.

Example of Professional Societies / Associations:

- *American Statistical Association (ASA)*
- *American Public Health Association (APHA)*
- *International Biometric Society (ENAR/IBS)*
- *Joint Statistical Meetings (JSM)*
- *Society for Clinical Trials (SCT)*

Dissertation

The PhD dissertation is expected to contain original results in statistical theory and methods in the solution of a problem which has relevance to a biomedical application. As a rule, the content of the dissertation should be adequate for publication in peer-refereed journals in the topic area of the dissertation. Students begin work on their dissertation research with the approval of their thesis sponsor and comprehensive examination committee. The only time limitation is the Graduate School of Arts and Sciences maximum of seven years from the time of enrollment in the doctoral program (the maximum is six years for those receiving advanced standing). Candidates who are making satisfactory progress toward finishing the dissertation have, upon application, been granted extensions by the Dean of GSAS, with the approval of their sponsor. With proper advising, PhD students should be able to finish the degree within five years of entry into the PhD program.

PhD candidates are required to submit an electronic copy of their final dissertation to the department. Electronic copies of past dissertations are available from the Director of Academic Programs.

DISSERTATION CONTINUED

In some cases a student may wish to change dissertation sponsors – for instance, if the student’s research leads to different areas of expertise than originally anticipated. In such cases, the student may seek approval from a new faculty sponsor. The candidate must inform the Doctoral Program Subcommittee Chair and the previous sponsor that the new sponsor will assume the previous sponsor’s duties. At this point, the student may also decide to pursue a new dissertation topic, with approval of the new sponsor, but in all cases the rules governing time limits and extensions still apply. Upon completion of the dissertation, and with approval of the candidate’s dissertation committee, the dissertation defense is scheduled.

PhD candidates are required to submit an electronic copy of their final dissertation to the department.

For more details regarding the PhD dissertation, the student is referred to the Dissertation Office website: www.gsas.columbia.edu/dissertations. The GSAS Dissertation Office is located on the Columbia Morningside Heights campus at 107 Low Memorial Library, 535 W. 116th Street, New York, NY 10027. Information is also available in the Department of Biostatistics and the Dean’s Office of GSAS on Morningside Campus.

Some Past PhD Dissertation Titles

The titles below are provided to give students some idea of topics that in past years have proved appropriate for the PhD degree:

- Optimal Treatment Regimes for Personalized Medicine and Mobile Health*, Eun Jeong Oh (2020)
- Statistical Learning Methods for Depression Screening and Intervention, and Structured Missing Imputation*, Huichen Zhu (2019)
- Quantile regression for zero-inflated outcomes*, Wodan Ling (2019)
- Functional Data Analytics for Wearable Device and Neuroscience Data*, Julia Wrobel (2019)
- Statistical Methods for Constructing Heterogenous Biomarker Networks*, Shanghong Xie (2019)
- Machine Learning Methods in Personalized Medicine Using Electronic Health Records*, Peng Wu (2019)
- Statistical Methods for Genetic Studies with Family History of Diseases*, Annie Lee (2019)
- Varying-Coefficient Models and Functional Data Analysis for Dynamic Network and Wearable Device Data*, Jihui Lee (2018)
- Statistical Methods for Modeling Biomarkers of Neuropsychiatric Disease*, Ming Sun (2018)
- Flexible Regression Models for Estimating Interaction between a Treatment and Scalar/Functional Predictors*, Hyung Park (2018)
- Statistical Learning Methods for Personalized Medicine*, Xin Qiu (2018)
- Methods in functional data analysis and functional genomics*, Daniel Backenroth (2018)
- Marginal screening on survival data*, Tzu-Jung Huang (2017)
- Methods for functional regression and nonlinear mixed-effect models with applications to PET data*, Yakuan Chen (2017)
- Statistical Learning Methods for Personalized Medical Decision Making*, Ying Liu (2016)
- Survival Analysis using Bivariate Archimedean Copulas*, Krishnendu Chandra (2015)
- Learning Logic Rules for Disease Classification: With an Application to Developing Criteria Sets for the Diagnostic and Statistical Manual of Mental Disorders*, Christine Mauro (2015)
- Empirical likelihood tests for stochastic ordering based on censored and biased data*, Hsin-wen Chang (2015)
- Sequential Designs for Individualized Dosing in Phase I Cancer Clinical Trials*, Xuezhou Mao (2015)
- Methods for Handling Measurement Error and Sources of Variation in Functional Data Models*, Xiaochen Cai (2015)
- Statistical modeling and statistical learning for disease clarification and prediction*, Tianle Chen (2014)
- Sequential Quantile Estimation Using Continuous Outcomes with Applications in Dose Finding*, Chih-Chi Hu (2014)
- Two-stage continual reassessment method and patient heterogeneity for dose-finding studies*, Xiaoyu Jia (2014)
- Wavelet-based Procedures for Scalar-on-function Regression*, Adam Ciarleglio (2013)
- Sample Size Calculations for the Semiparametric Analysis of Short-term and Long-term Hazard Ratios*, Yi Wang (2013)
- On Composition Data Modeling and Its Biomedical Applications*, Bingzhi Zhang (2013)

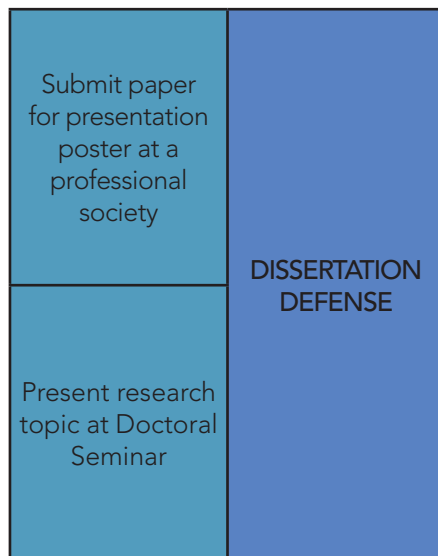


Typical PhD Timeline (PhD student entering with a relevant Masters)

Fall I	Spring I	Summer I	Fall II	Spring II			
P9104 Probability for Biostatisticians		QUALIFYING EXAMS Written/Take-home	P9111 Asymptotics	Formalize Research Topic	Submit paper for presentation poster at a professional society	DISSERTATION DEFENSE	
	P8160 Topics in Advanced Statistical Computing		P9120 Topics in Statistical Learning & Data Mining I				Present research topic at Doctoral Seminar
P9109 Theory of Statistical Inference I	P9110 Theory of Statistical Inference II						
P9130 Advanced Methods I Electives	P9185 SPRIS I Electives			P9186 SPRIS II Electives	Prepare for Oral Exam		
Statistical Inference Problem Seminar							
Attend all Departmental Colloquium and Research Talks							

Typical PhD Timeline (PhD student entering without a relevant Masters)

Fall I	Spring I	Summer I	Fall II	Spring II	Summer II	Fall III	Spring III
P6400 Principles of Epidemiology	P8109 Statistical Inference	Review material; begin research projects	P9104 Probability for Biostatisticians	P8160 Topics in Advanced Statistical Computing	QUALIFYING EXAMS (Written/Take-Home)	P9111 Asymptotics	Formalize Research Topic
P8104 Probability	P8131 Biostatistical Methods II		P9109 Theory of Statistical Inference I			P9110 Theory of Statistical Inference II	
P8105 Data Science	P8106 Data Science II Elective	P9130 Advanced Methods I	P9185 SPRIS I	P9186 SPRIS II Electives		Prepare for Oral Exam	
P8130 Biostatistical Methods I							Statistical Inference Problem Seminar
Attend all Departmental Colloquium and Research Talks							



****Attend all department colloquia and research talks****



Certificate Program

Certification of Professional Achievement in Clinical and Translational Research

(CPA-CTR) provides effective training in core methods used in clinical and translational research in a condensed format suited to working clinicians and post-doctoral researchers. The 12-credit program may be completed in as little as six months; ten credits are earned over 5-weeks in the summer through the Columbia Summer Research Institute (CSRI), and the remaining two credits are earned in a subsequent Fall or Spring semester. In this format, working professionals are able to receive graduate level training in research methods while simultaneously working with a multi-disciplinary mentoring team to develop a research project.

The Certification of Professional Achievement in Clinical and Translational Research curriculum is comprised of intensive instruction in biostatistics and epidemiology, followed by more specialized coursework, and culminates in a full NIH-style grant proposal or an original paper to be submitted to a peer-reviewed journal. The selected coursework and key deliverables emphasize quantitative training, critical thinking skills, and practical strategies in order for junior investigators to be competitive in the quest for independent grant funding. CPA-CTR requirements must be completed within two years of enrollment in CSRI.

Program Requirements

1) Successful completion of the CSRI, a 10-credit summer program consisting of five required courses:

- P6104 Introduction to Biostatistical Methods
- P6400 Principles of Epidemiology
- P8182 Writing a Successful NIH Grant
- P8568 Decision Analysis for Public Health and Clinical Practices
- P8750 Race and Health

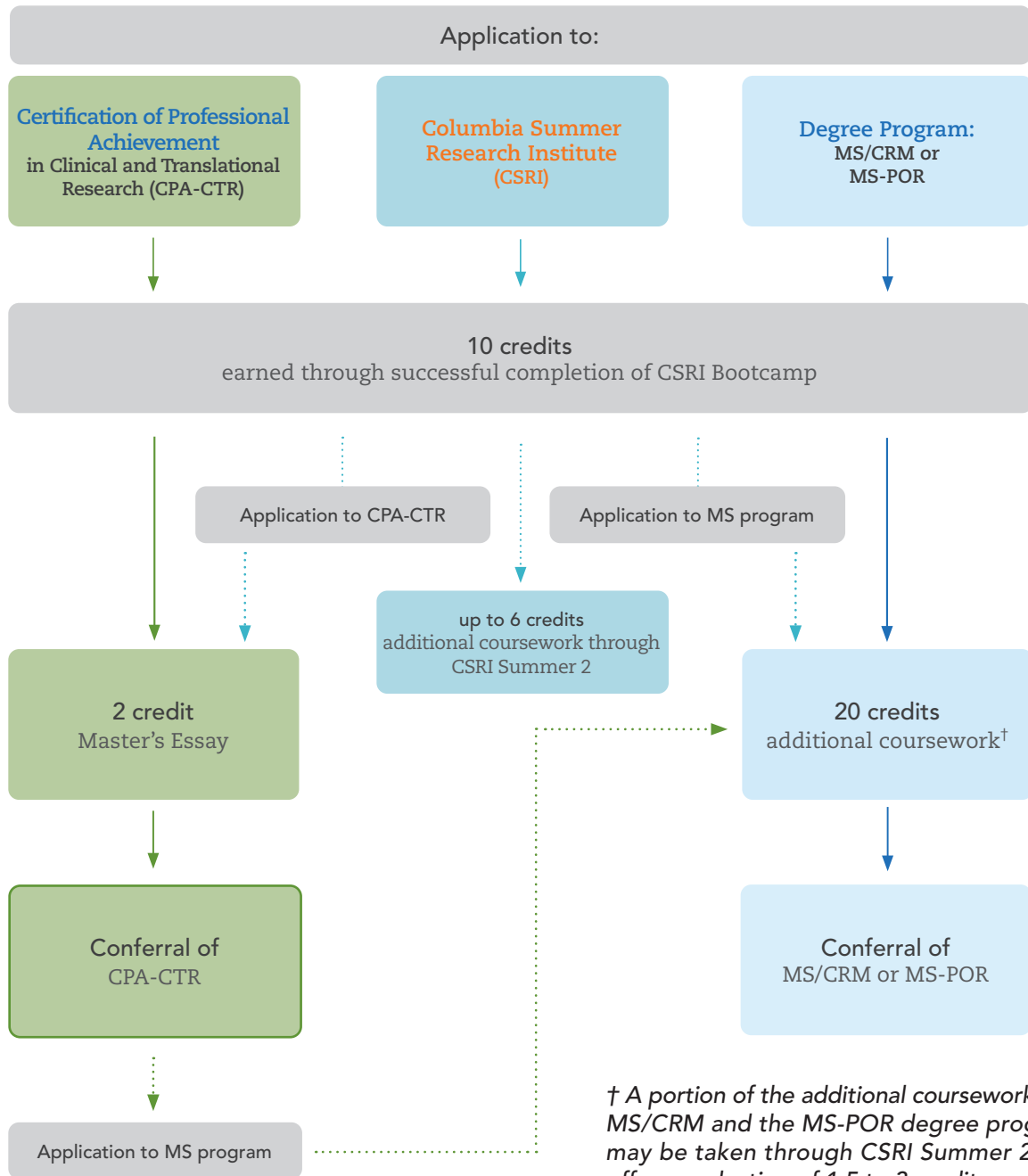
2) Successful completion of a 2-credit Master's Essay course - select one:

- P9160 Master's Essay (original research article) - Spring
- P9165 Master's Essay (full grant proposal) - Fall

Key Deliverables

- NIH-style grant outline for either an R01 or K award (all)
- Full grant proposal for an R01 or K award (P9165)
- Original research article submitted to a peer-reviewed journal (P9160)
- Critical thinking skills
- Foundational analytic toolbox
- The application process for the CPA-CTR is the same as the process for CSRI Bootcamp. CSRI alums who wish to earn the CPA-CTR should contact csri@columbia.edu for next steps.

Research Training Options



† A portion of the additional coursework for the MS/CRM and the MS-POR degree programs may be taken through CSRI Summer 2 which offers a selection of 1.5 to 3-credit courses that run from 1 week to 6 weeks in length.

Department of Biostatistics Courses

These are the courses offered by the Department of Biostatistics. Due to faculty commitments, the frequency of the courses changes from time to time. Students are advised to check the current schedule of courses listed on the MSPH web page: www.mailman.columbia.edu/academics/courses. Students may also review the course offerings of the Statistics Department at the Morningside Campus in the Graduate School of Arts and Sciences: www.stat.columbia.edu.

Students are encouraged to meet with their faculty advisors at least twice a year – in the fall and in the spring. Permission is not required for approved courses in a student's approved program of study. Students must first obtain permission from their faculty advisors to take courses outside the approved program. **Failure to comply with these guidelines may jeopardize plans for graduation.**

P6103 Introduction to Biostatistics 3 points

Prerequisites: Permission of the instructor required for all non-Public Health students.

Biostatistics is essential to ensuring that findings and practices in public health and biomedicine are supported by reliable evidence. This course covers the basic tools for the collection, analysis, and presentation of data in all areas of public health. Central to these skills is assessing the impact of chance and variability on the interpretation of research findings and subsequent recommendations for public health practice and policy. Topics covered include: general principles of study design; hypothesis testing; review of methods for comparison of discrete and continuous data including ANOVA, t-test, correlation, and regression.

P6104 Introduction to Biostatistical Methods 3 points

Prerequisites: Placement exam required, and the instructor's permission.

An enriched core course for students concentrating in biostatistics and others who expect to take additional courses in biostatistics beyond the two main second-level courses (P8100 and P8120). It covers in greater depth all of the topics in P6103 and is the best preparation for students anticipating a quantitative orientation in their degree programs. Topics covered include standard distributions, measures of central tendency and dispersion, hypothesis testing, point estimation, confidence intervals, and an introduction to correlation and regression.

P6107 Biostatistical Analysis Using SPSS 3 points

Prerequisites: P6103 or P6104 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

A logical follow-up course to an introductory biostatistics course. Covers uses of the computer in cleaning, summarizing, and cross-classifying data. Enhancement of the material covered in P6103/P6104— including regression, correlation, and contingency table analysis, and the analysis of variance— with data analysis carried out using SAS software.

P6110 Statistical Computing with SAS 3 points

Prerequisites: P6103 or P6104 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

A logical follow-up course to an introductory biostatistics course. Covers uses of the computer in cleaning, summarizing, and cross-classifying data. Enhancement of the material covered in P6103/P6104— including regression, correlation, and contingency table analysis, and the analysis of variance— with data analysis carried out using SAS software.

P6170 New Drug Development: A Regulatory Overview 3 points

Prerequisite: P6103 or P6104, and P6400.

Provides our CTSA fellows and scholars with insights into and understanding of the process of patient oriented/translational research and gives them an opportunity to meet active investigators from academia and industry, and learn about some career enhancing resources available at CUMC. Active researchers from various clinical disciplines and public health are invited to speak on research techniques, design, and laboratory methodology as applied to current studies. They present their experiences in conducting patient orientated research on the Health Sciences campus and elsewhere. Also features speakers from both the pharmaceutical and biotech industries who discuss drug development, and preclinical and clinical trials. Other lectures deal with FDA regulations, patent law, and the Institutional Review Board and ways to effectively build and succeed in a clinical/translational academic career.

P8100 Applied Regression I 3 points

Prerequisites: P6103 or P6104 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

This course will provide an introduction to the basics of regression analysis. The class will proceed systematically from the examination of the distributional qualities of the measures of interest, to assessing the appropriateness of the assumption of linearity, to issues related to variable inclusion, model fit, interpretation, and regression diagnostics.

P8101 Introduction to Health Data Science 3 points

Prerequisites: P6103 or P6104 or MPH Quantitative Foundations

This course will introduce students to core data science skills and concepts through the exploration of applied biostatistics. The course will begin with an introduction to the R programming language and the RStudio IDE, focusing on contemporary tidyverse functions and reproducible programming methods. Then, the course will instruct students in contemporary data manipulation and visualization tools while systematically covering core applied biostatistics topics, including confidence intervals, hypothesis testing, permutation tests, and logistic and linear regression. Finally, the semester will end with an introduction to machine learning concepts, including terminology, best practices in test/training sets, cross-validation, and a survey of contemporary classification and regression algorithms.

P8103 Colloquium on Patient Oriented Research 2 points (0.5 points x 4 semesters)

Prerequisite: This is a required course for MS-POR students, CTSA K12 scholars, and CTSA T32 trainees. Permission required.

Provides our CTSA fellows and scholars with insights into and understanding of the process of patient oriented/translational research and gives them an opportunity to meet active investigators from academia and industry, and learn about some career enhancing resources available at CUMC. Active researchers from various clinical disciplines and public health are invited to speak on research techniques, design, and laboratory methodology as applied to current studies. They present their experiences in conducting patient orientated research on the Health Sciences campus and elsewhere. Also features speakers from both the pharmaceutical and biotech industries who discuss drug development, and preclinical and clinical trials. Other lectures deal with FDA regulations, patent law, and the Institutional Review Board and ways to effectively build and succeed in a clinical/translational academic career.

P8104 Probability 3 points

Prerequisites: P6104 or P8130 (may be corequisite), working knowledge of calculus, and the instructor's permission.

Topics include: Fundamentals, random variables, and distribution functions in one or more dimensions; moments, conditional probabilities, and densities; Laplace transforms and characteristic functions.

Infinite sequences of random variables, weak and strong large numbers; central limit theorem.

P8105 Data Science I 3 points

Prerequisites: Experience in R programming (or programming in another language) and data analysis is recommended and the instructor's permission.

Contemporary biostatistics and data analysis depends on the mastery of tools for computation, visualization, dissemination, and reproducibility in addition to proficiency in traditional statistical techniques. The goal of this course is to provide training in the elements of a complete pipeline for data analysis.

P8106 Data Science II 3 points

Prerequisites: P8105 Data Science I

With the explosion of "Big Data" problems, statistical learning has become a very hot field in many scientific areas. The goal of this course is to provide the training in practical statistical learning. It is targeted to MS students with some data analysis experience.

P8107 Introduction to Mathematical Statistics 3 points

Prerequisites: MPH Quantitative foundations or P6104 (Not open to Theory & Methods, Statistical Genetics, or Pharmaceutical Statistics tracks)

The first portion of this course provides an introductory-level mathematical treatment of the fundamental principles of probability theory, providing the foundations for statistical inference. Students will learn how to apply these principles to solve a range of applications. The second portion of this course provides a mathematical treatment of (a) point estimation, including evaluation of estimators and methods of estimation; (b) interval estimation; and (c) hypothesis testing, including power calculations and likelihood ratio testing.

P8108 Survival Analysis 3 points

Prerequisites: P8104, P8109, P8130, and the instructor's permission.

This course focuses on methods for the analysis of survival data, or time-to-event data. Survival analysis is a method for survival data or failure (death) time data, that is time-to-event data, which arises in a number of applied fields, such as medicine, biology, public health, epidemiology, engineering, economics, and demography. A special course of difficulty in the analysis of survival data is the possibility that some individual may not be observed for the full time to failure. Instead of knowing the failure time t , all we know about these individuals is that their time-to-failure exceeds some value y where y is the follow-up time of these individuals in the study. Students in this class will learn how to make inference for the event times with censored.

P8109 Statistical Inference 3 points

Prerequisites: P8104, working knowledge of calculus and linear algebra, and the instructor's permission.

This course covers a review of mathematical statistics and probability theory at the Masters level. Students will be exposed to theory of estimation and hypothesis testing, confidence intervals and Bayesian inference. Topics include population parameters, sufficient statistics, basic distribution theory, point and interval estimation, introduction to the theory of hypothesis testing, and nonparametric procedures.

P8110 Applied Regression II 3 points

Prerequisites: P6103 or P6104 or MPH Quantitative Foundations, Evidence, and Policy core course, P8100, and the instructor's permission.

An introduction to the application of statistical methods in survival analysis, generalized linear models, and design of experiments. Topics to be covered include estimation and comparison of survival curves, regression models for survival data, log-linear models, logit models, analysis of repeated measurements, and the analysis of data from blocked and split-plot experiments. Examples are drawn from the health sciences.

P8112 Systematic Review and Meta-Analysis 1.5 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, and P6400.

Research synthesis using systematic review and meta-analysis is one of the most valuable of research endeavors, and can be a particularly rewarding experience for junior investigators who want to develop expertise in a specific area of public health or medicine by producing a product with significant scientific impact. This course will combine lecture and workshop elements to introduce students to the principles and practices of systematic review and meta-analysis. It will be targeted to students who have previously been introduced to the concepts of basic biostatistics, epidemiology, and clinical trials. By the end of the course, each student (as part of a group) will have conducted a systematic review and meta-analysis of a topic chosen by their group. The emphasis in this course is on the actual conduct of systematic review and meta-analysis, so morning lectures are followed daily by afternoon lab/workshop hours during which instructors support students in mastering each stage of the research process. As a culminating experience, students make oral presentations summarizing the process they followed, unexpected challenges, main findings, and plans for future research using their newly acquired research synthesis skills.

P8116 Design of Medical Experiments 3 points

Prerequisites: P8104, P8109, P8130, and the instructor's permission.

This course covers the fundamental principles and techniques of experimental designs in clinical studies. Topics include reliability of measurement, linear regression analysis, parallel groups design, analysis of variance (ANOVA), multiple comparison, blocking, stratification, analysis of covariance (ANCOVA), repeated measures studies; Latin squares design, crossover study, randomized incomplete block design, and factorial design.

P8119 Advanced Statistical and Computational Methods in Genetics and Genomics 3 points

Prerequisites: P6104 or P8130, and the instructor's permission.

This course introduces students to advanced computational and statistical methods used in the design and analysis of high-dimensional genetic data, an area of critical importance in the current era of Big Data. The course starts with a brief background in genetics, followed by in depth discussion of topics in genome-wide linkage and association studies, and next-generation sequencing studies. Additional topics such as network genetics will also be covered. Examples from recent and ongoing applications to complex traits will be used to illustrate methods and concepts.

P8120 Analysis of Categorical Data 3 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, P6400 (may be corequisite), and the instructor's permission.

A comprehensive overview of methods of analysis for binary and other discrete response data, with applications to epidemiological and clinical studies. Topics discussed include the fourfold table, significance versus magnitude of association; estimation of relative risk; matching in design and analysis; interrater agreement; logistic regression analysis.

P8122 Statistical Methods for Causal Inference 3 points

Prerequisites: P8100 and P8110 or P8130 and P8131

Substantive questions in empirical scientific and policy research are often causal. This class will introduce students to both statistical theory and practice of causal inference. As theoretical frameworks, we will discuss potential outcomes, causal graphs, randomization and model-based inference, causal mediation, and sufficient component causes. We will cover various methodological tools including randomized experiments, matching, inverse probability weighting, instrumental variable approaches, dynamic causal models, sensitivity analysis, statistical methods for mediation and interaction. We will analyze the strengths and weaknesses of these methods. The course will draw upon examples from social sciences, public health, and other disciplines. The instructor will illustrate application of the approaches using R/ SAS/STATA software. Students will be evaluated and will deepen the understanding of the statistical principles underlying the approaches as well as their application in homework assignments, a take home midterm, and final take home practicum.

P8123 Analysis of Health Surveys 3 points

Pre-requisites: P8131 (or P8110) and P8104 (or P8107)

This is an applied statistical methods course. The course will introduce main techniques used in sampling practice, including simple random sampling, stratification, systematic sampling, cluster sampling, probability proportional to size sampling, and multistage sampling. Using national health surveys as examples, the course will introduce and demonstrate the application of statistical methods in analysing across-sectional surveys and repeated and longitudinal surveys, and conducting multiple imputation for missing data in large surveys. Other topics will include methods for variance estimation, weighting, post-stratification, and non-sampling errors. If time allows, new developments in small area estimation and in the era of data science will also be discussed.

P8130 Biostatistical Methods I 3 points

Prerequisites: Students are required to have working knowledge of calculus and linear algebra. Some programming experience with R/R Studio is recommended, but not required.

This one-semester course introduces basic applied descriptive and inferential statistics. The first part of the course includes elementary probability theory, an introduction to statistical distributions, principles of estimation and hypothesis testing, methods for comparison of discrete and continuous data including chi-squared test of independence, t-test, analysis of variance (ANOVA), and their non-parametric equivalents. The second part of the course focuses on linear models (regression) theory and their practical implementation.

P8131 Biostatistical Methods II 3 points

Prerequisites: P8130 and instructor's permission for non-MS Biostatistics students.

Regression analysis is widely used in biomedical research. Non-continuous (e.g., binary or count-valued) responses, correlated observations, and censored data are frequently encountered in regression analysis. This course will introduce advanced statistical methods to address these practical problems. Topics include generalized linear models (GLM) for non-Gaussian response, mixed-effects models and generalized estimating equations (GEE) for correlated observations, and Cox proportional hazards models for survival data analysis. Examples are drawn from biomedical sciences.

P8133 Bayesian Analysis and Adaptive Designs for Clinical Trials 3 points

Prerequisites: P8104 and P8109 or their equivalents, and the instructor's permission.

An introduction to sequential analysis as it applies to statistical problems in clinical trials, hypothesis testing, selection, and estimation. Emphasis is placed on a study of procedures, operating characteristics, and problems of implementation, rather than mathematical theory. Students obtain an overview of currently available sequential designs and the advantages and disadvantages they offer in comparison with classical designs.

P8134 Stochastic Approximation and Modern Dose-Finding 3 points

Prerequisites: P8104 and P8109 or their equivalents, and the instructor's permission.

Provides an in-depth study of statistical designs for dose-finding clinical trials of new drugs. This course is designed for advanced Master's, DrPH, and PhD students in biostatistics. The overall learning objective is to equip students with the techniques to construct, evaluate, and critique dose-finding designs. The course consists of two parts. The first is a review of modern dose-finding techniques with a focus on the continual reassessment method (CRM) and its clinical applications. The second part presents advanced topics on stochastic approximation and its related theory. Connections between the dose-finding methods (part 1) and the stochastic approximation (part 2) will be drawn. The practical implication of these connections is two-fold. First, the stochastic approximation will provide a versatile and mathematically rigorous framework for tailoring dose-finding designs to specific clinical situations. Second, the well-studied theory of stochastic approximation will be an effective analytical tool to approximate the theoretical properties of the CRM.

P8139 Statistical Genetics Modeling 3 points

Prerequisites: P6103 or P6104 or P8130, a working knowledge of calculus, and the instructor's permission.

Present to students statistical tools so that they can grasp the fundamentals of the design, conduct and analysis of genetic association studies. The course will thoroughly discuss current methods that are being used to map genes for common complex diseases. Great emphasis will be placed on candidate-gene and genome-wide association studies, but linkage methods will also be treated. Another key feature of this course will be a detailed treatment of the major findings of the Human Genome Project and HapMap Project.

P8140 Introduction to Randomized Clinical Trials 3 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

Fundamental methods and concepts of the randomized clinical trial: protocol development, randomization, blindedness, patient recruitment, informed consent, compliance, sample size determination, crossovers, collaborative trials. Each student prepares and submits the protocol for a real or hypothetical clinical trial.

P8142 Clinical Trial Methodology 3 points

Prerequisites: P6103 or P6104 or P8130, and the instructor's permission.

The main objective of this course is to provide students and investigators with a working knowledge of certain methodological issues that arise in designing a clinical trial in order to conduct complex study designs that yield valid and reliable results. With emphasis on several methodological and practical issues related to the design and analysis of clinical experiments, topics include: the design of small studies (Phase I and II studies), interim analyses and group sequential methods, survival studies, multiple outcome measures, surrogate outcomes, multicenter studies, issues in data analysis, and reporting and interpreting study results.

P8144 Pharmaceutical Statistics 3 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission. SAS knowledge recommended.

Drug development from compound discovery to marketing and commercialization registration is a lengthy and complex process in which statisticians play an important role from beginning to end. The main objective of this course is to provide students with working knowledge of the methodological and operational issues that arise in different stages of drug development that involve statistical contributions. Topics include: Introduction of drug development; design and analysis of toxicology studies, pharmacokinetics and pharmacodynamics studies, biomarker studies, Phase I/II/III studies; issues in clinical studies such as non-inferiority studies, meta-analysis, and endpoint selection; and preparation for the FDA advisory committee drug approval process. In addition, the views and positions of different regulatory bodies, such as the FDA or EMEA, on design and analysis will be discussed.

P8149 Human Population Genetics 3 points

Prerequisites: P8104 and the instructor's permission.

This course will cover all statistical aspects of population genetics. Upon completion of this course, the students will be able to model and do inference of underlying population genetic mechanisms and apply acquired knowledge about population genetics to the analyses of phenotypes.

P8157 Analysis of Longitudinal Data 3 points

Prerequisites: P8104, P8109, P8130, and the instructor's permission.

The course will introduce students to statistical models and methods for longitudinal data, i.e., repeatedly measured data over time or under different conditions. The topics will include design and sample size calculation, Hotelling's T^2 , multivariate analysis of variance, multivariate linear regression (generalized linear models), models for correlation, unbalanced repeated measurements, mixed effects models, EM algorithm, methods for non-normally distributed data, generalized estimating equations, generalized linear mixed models, and missing data.

P8158 Latent Variable and Structural Equation Modeling for Health Sciences 3 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

This course is designed for those students (or any researchers) who want to gain a significant familiarity with a collection of statistical techniques that target the measurement of latent variables (i.e. variables that cannot be measured directly) as well as methods for estimating relationships among variables within causal systems. This course covers: both continuous and categorical latent variable measurement models (i.e. exploratory and confirmatory factor analysis, item response theory models, latent class and finite mixture models), as well as estimation of relationships in hypothesized causal systems using structural equation modeling. Data analysis examples will come from health science applications and practical implementation of all methods will be demonstrated using predominately the Mplus software, but also the R software.

P8160 Topics in Advanced Statistical Computing 3 points

Prerequisites: P8109, a basic understanding of Bayesian inference, working knowledge of a programming language, and the instructor's permission.

As statistical models become increasingly complex, it is often the case that exact or even asymptotic distributions of estimators and test statistics are intractable. With the continuing improvement of processor speed, computationally intensive methods have become invaluable tools for statisticians to use in practice. This course covers the basic modern statistical computing techniques and how they are applied in a variety of practical situations. Topics include numerical optimization, random number generation, simulation, Monte Carlo integration, permutation tests, jackknife and bootstrap procedures, Markov Chain Monte Carlo methods in Bayesian settings, and the EM algorithm.

P8180 Relational Databases and SQL Programming for Research and Data Science 3 points

Prerequisites: P6103 or P6104 or P8130 or MPH Quantitative Foundations, Evidence, and Policy core course, and the instructor's permission.

This class provides an overview of the specific techniques available to collect, store, retrieve, and control the quality of data in research projects. Students will be introduced to these concepts through a combination of lecture videos and a substantial hands-on component consisting of structured computer-based exercises. Spreadsheet and database technologies will be reviewed in detail to establish guidelines as to the appropriateness of their use to manage data in research. Microsoft Excel will be used to demonstrate the limitations of the spreadsheet as a tool in data management. Students will use Microsoft Access to learn the principles, concepts, and powerful capabilities of relational databases. Basics of the SQL language will be covered to teach students how to retrieve information from a database, and to create and fill database tables. SAS and MySQL software will be covered briefly to give students exposure to performing data management tasks in these commonly-used programs.

P8182 Writing a Successful Grant Application 1 points

Prerequisites: Concurrent enrollment in the Columbia Summer Research Institute. Required for MS-POR students.

This seminar-style course will lead students through the process of writing an NIH-style grant application. By the end of the course, each student submits a research proposal outline following NIH guidelines for either an R01 or K (career development) award. The emphasis in this course is on the quality of the proposed research, taking into account feasibility, relevance, innovation, ethical foundation, and public health impact. As a culminating experience, students make oral presentations summarizing their research proposals to an invited panel of senior, experienced CUMC faculty, and receive feedback on their proposed research aims and approaches.

P8185 Capstone Consulting Seminar 1 point

Prerequisites: At least 15 points of required coursework in biostatistics, and the instructor's permission.

Required capstone course for the MS/PS, MS/SG, MS/TM and 2-year MPH students in Biostatistics. Provides experience in the art of consulting and in the proper application of statistical techniques to public health and medical research problems. Enables students to translate research objectives into statistical hypotheses, devise appropriate study designs, perform sample size calculations for studies employing simple random sampling, formulate and prepare written plans for statistical analysis for a research proposal, compose summaries of quantitative analyses, and communicate results clearly to public health colleagues. Based on seminars requiring active student participation.

P9104 Probability for Biostatisticians 3 points

Pre-requisites: P8109 and P8110, advanced calculus. Instructor's permission for MS students

The biostatistical field is changing with new directions emerging constantly. Doing research in these new directions, which often involve large data and complex designs, requires advanced probability and statistics tools. The purpose of this new course is to collect these important probability methods and present them in a way that is friendly to a biostatistics audience. This course is designed for PhD students in Biostatistics. Its primary objective is to help the students achieve a solid understanding of these probability methods and develop strong analytical skills that are necessary for conducting methodological research in modern biostatistics. At the completion of this course, the students will a) have a working knowledge in Law of Large Numbers, Central Limit Theorems, martingale theory, Brownian motions, weak convergence, empirical process, and Markov chain theory; b) be able to understand the biostatistical literature that involves such methods; c) be able to do proofs that call for such knowledge.

P9109 Theory of Statistical Inference I 3 points

Prerequisites: P8104, P8109, and the instructor's permission.

This course offers a general introduction to essential materials in advanced statistical theory for doctoral students in biostatistics. The course is designed to prepare doctoral students in biostatistics for their written theory qualifying exam. Students in this course will learn theory of estimation, confidence sets and hypothesis testing. Specific topics include a quick review of measure-theoretic probability theory, concepts of sufficiency and completeness, unbiased estimation (UMVUE), least squares principle, likelihood estimation, a variety of estimators and their asymptotic properties, confidence sets, the Neyman-Pearson lemma and uniformly most powerful tests. If time permits, the likelihood ratio test, score test and Wald test, and sequential analysis will be covered.

P9110 Theory of Statistical Inference II 3 points

Prerequisites: P8104, P8109, P9109, and the instructor's permission. This is a continuation of P9109.

This course continues the introduction to mathematical statistics for doctoral students in biostatistics. Topics to be covered include: principles of decision theory, Bayesian estimation, Hypothesis testing, asymptotics, M-estimation, Wald tests, and score tests.

P9111 Asymptotic Statistics 3 points

Prerequisites: P8104, P8109, P9109, P9110, and the instructor's permission.

The choice of topics will vary from year to year, but will typically include: empirical processes and M-estimation, bootstrap methods, empirical likelihood, contiguity, local asymptotic normality, counting process methods in survival analysis, semiparametric inference and efficiency.

P9120 Topics in Statistical Learning and Data Mining I 3 points

Prerequisites: Intended for Biostatistics PhD students in their second year or higher and theoretically inclined DrPH or MS students of Biostatistics.

Provide students a systematic training in key topics in modern supervised statistical learning and data mining. For the most part, the focus will remain on a theoretically sound understanding of the methods (learning algorithms) and their applications in complex data analysis, rather than proving technical theorems. Applications of the statistical learning and data mining tools in biomedical and health sciences will be highlighted.

P9130 Advanced Biostatistical Methods I 3 points

Prerequisites: Advanced calculus, linear algebra, basic probability, statistical inference and instructor's permission for MS students

The course will provide a solid foundation of the theory behind linear models and generalized linear models. More emphasis will be placed on concepts and theory with mathematical rigor. Topics covered including linear regression models, logistic regression models, generalized linear regression models and methods for the analysis contingency tables.

P9160 Master's Essay in Biostatistics: Clinical Research Methods 3 points

Prerequisites: At least 15 points of required coursework, and the instructor's permission.

Attention: This course is for MS/CRM students only.

Students produce a Master's essay in the form of a research article of publishable quality, supervised by faculty members from Biostatistics and from the student's own clinical field.

P9165 Master's Essay in Biostatistics: Patient Oriented Research 0 points

Prerequisites: At least 15 points of required coursework, and the instructor's permission.

Attention: This course is for MS-POR students only.

Students produce a Master's essay in the form of an NIH-style grant application, supervised by a project sponsor from Biostatistics and a mentor from the student's own clinical field. A formal presentation to the POR advisory board is required for successful completion of the course.

P9185 Statistical Practices and Research for Interdisciplinary Sciences (SPRIS) 3 points

Prerequisites: At least 15 points of required coursework, and the instructor's permission.

Required course for the DrPH and PhD students in biostatistics. Provides experience in the art of consulting and in the proper application of statistical techniques to public health and medical research problems. Enables students to translate research objectives into statistical hypotheses, devise appropriate study designs, perform sample size calculations for studies employing simple random sampling, formulate and prepare written plans for statistical analysis for a research proposal, compose summaries of quantitative analyses, and communicate results clearly to public health colleagues. Based on seminars requiring active student participation.

P9186 Statistical Practices and Research for Interdisciplinary Science (SPRIS) II 1.5 points

Prerequisites: P9185

Students will apply the concepts and methods introduced in Statistical Practices and Research for Interdisciplinary Science (SPRIS) I to a real research setting. Each student will be paired with a Biostatistics faculty member. The student will participate in one of the mentor's collaborative projects to learn how to be an effective member of an interdisciplinary team. Student experience will vary depending on the assigned faculty member, but all students will gain exposure to preparing collaborative grant applications, designing research studies, analyzing real data, interpreting and presenting results, and writing manuscripts. Mentors will help to develop the student's data intuition skills, ability to ask good research questions, and leadership qualities. Where necessary, students may replicate projects already completed by the faculty mentor to gain experience.

P6190, P8190, P9190 Tutorials in Biostatistics 1 to 6 points

For appropriately qualified students wishing to enrich their programs by undertaking literature reviews, special studies, or small group instruction in topics not covered in formal courses. Hours to be arranged.

89260 Building Interdisciplinary Research Models 2 points

Interdisciplinary research is an approach to advancing scientific knowledge requiring mastery of specific competencies. This seminar will introduce the students to competencies in interdisciplinary research through a combination of readings and lectures in each necessary aspect, chosen from fields essential to successful interdisciplinary research. This course will assist learners to understand why and how different professional disciplines, each representing a body of scientific knowledge, must work together to generate and disseminate knowledge. Learners will develop a set of skills specific to be an effective member and leader of an interdisciplinary research team, and will become familiar with the advantages of team science.

Colloquia

During the Fall and Spring semesters, the Department of Biostatistics holds seminars on a wide variety of topics which are of interest to both students and faculty. The speakers are occasionally departmental faculty members themselves but very often are invited guests who spend the day of their seminar discussing their research with Biostatistics faculty and students. While all students are strongly urged to attend, **doctoral student attendance is mandatory.**

Consulting Service

All MS/PS, MS/SG, MS/TM, MPH are required to participate in the Biostatistics Consulting Service. This program is designed to enable students to demonstrate their ability to integrate academic studies with the role of biostatistical consultant/collaborator. The Biostatistics Consulting Service offers advice on data analysis and appropriate methods of data presentation for publications, and provides design recommendations for public health and clinical research, including preparation of grant proposals.

Participation in the Biostatistics Consulting Service meets the MS/PS, MS/SG, MS/TM and MPH capstone requirement while providing students with an opportunity to gain invaluable experience working with a diverse clientele on a variety of statistical problems.

Teaching Assistantships

Each semester, the Department makes available a limited number of Teaching Assistant (TA) positions. Upon completion of one full semester of course work, eligible students may apply for a TA slot. Students are advised to carefully consult the following policy on qualification, selection, and compensation of TAs before considering one of these positions. All TA candidates must apply to the Assistant Director of Academic Programs.

To qualify for a TA position, students must:

- be registered as a full-time student during the the semester of the TA opportunity
- NOT be employed by another department at Columbia University
- have successfully completed the course of interest
- maintain a GPA of 3.3 or better
- be able to devote several hours per week to TA duties. This includes, but is not limited to:
 - Attending class for lectures
 - Recitation periods (for core teaching assistants)
 - 1-2 regularly scheduled office hours
 - Homework grading and preparation of teaching materials

Selection of TAs is made by the instructor. Priority is given to students in doctoral programs, students with greater seniority, and students with previous TA experience who have received good evaluations from their former students and course instructors.

TA compensation is taxable and is paid out over the course of the semester.



The Practicum Requirement

The intent of the practicum requirement is to engage students in activities aligned with their career goals, as well as activities that demonstrate application of biostatistical methods and public health concepts relevant to the student's area of interest. Students will seek out activities that further develop their skill set and add new tools to their professional toolkit. Upon completion of the practicum, the student will be able to provide evidence of application of these skills to potential employers. Practicum placements are made on an individual basis in consultation with faculty advisors who must approve both the proposed practicum prior to its initiation. Students and mentors must complete an evaluation at the conclusion of the practicum experience.

Goals of the Practicum are:

For the Department of Biostatistics

- To provide the University with a part of the formative assessment of the student's ability to function as a Biostatistician; and
- To serve as a means of continually evaluating the relevance and effectiveness of the curriculum, leading to modifications of the formative and summative assessments when necessary.

For the Student

- To provide a continuing series of practical experiences geared to his or her level of expertise, which will offer a chance to apply principles, skills, and techniques that have been acquired;
- To help the student learn how to assume professional roles in a variety of practice settings while becoming accustomed to a range of organizational structures, working relationships, and job expectations; and
- To help the student develop professional identification as a Biostatistician and gain experience in fulfilling his or her role as a team member working with other professionals.

For the Mentoring Institution/Organization

- To provide mentorship input into the university program and, thereby, allow staff to share in the development of future Biostatisticians;
- To serve as a growth experience for the mentor's staff through interaction with the students; and
- To provide the mentor an opportunity to recruit employees and reduce the time needed for on-the-job training of any students who, upon graduation, are hired.

Practicum Roles

The **student** is responsible for identifying potential practicum sites and making arrangements for his or her practicum experience at an appropriate site. Appropriate sites will offer professional training and specialization. The practicum project can include but not limited to:

- Statistical method development
- Simulation study to compare existing methods
- Novel application of existing statistical methods
- Collaborative research that can lead to significant scientific findings

Sites must be approved by the student's faculty advisor. Ideally, the practicum placement should be approved no less than a month before the beginning of the practicum.

In addition to the student, there are three individuals with roles in the practicum experience, they are:

1. **Faculty Advisor:** the student's assigned advisor reviews and approves the proposed practicum as being relevant to the student's program track and career interests. The advisor must also endorse the student's suitability as well as his or her cognitive ability for a given practicum experience.
2. **Department Practicum Coordinator:** Corey Adams, (ca2594@cumc.columbia.edu). The Department Practicum Coordinator is responsible for supervision of the practicum experience once the advisor and mentor have approved the general concept and basic objectives for a given practicum.
3. **Practicum Mentor:** the field supervisor who provides the educational experience and mentorship, which are at the heart of the practicum experience. Practicum mentors should be motivated to host practicum experiences from a sense of professional commitment to help students achieve professional skills and status. For the purposes of the practicum, a qualified mentor may include database administrators, researchers, professors, doctors, etc. It is necessary, that a mentor fully operate effectively at a professional level in his/her field.

Student Requirements for Completing the Practicum Requirements

1. The student is responsible for finding a practicum and for completing the online **Practicum Approval Form**, at least one month before the start of the practicum. The form should be submitted no later than December 1st of the student's second year. **Two-year MPH students** must also complete a **Scope of Work (SOW)** form with the Office of Field Practice as well as the department's practicum form.
2. Once approved, the student begins his or her practicum experience. Before completing the practicum, the student and his or her advisor should meet at the midpoint to discuss progress.
3. After the completion of the practicum, the student submits a practicum report and a poster to his or her advisor for approval. The student will present his or her poster at the Annual Practicum Poster Symposium.
4. Links to the appropriate forms can be retrieved from Corey Adams, ca2594@cumc.columbia.edu

NOTE TO INTERNATIONAL STUDENTS

CPT and OPT forms must be obtained from ISSO and submitted to the Department Practicum Coordinator. Keep in mind that CPT is only authorized for dates that are within a semester. If you have already completed a CPT but have another opportunity that is different from your completed practicum, then you must register for PUBH 8086 – Public Health Practice Seminar (0.5 cr).

The Process of Completing the Practicum Requirement

