Milken Institute School of Public Health
THE GEORGE WASHINGTON UNIVERSITY

Department of Biostatistics and Bioinformatics

PhD in Health Data Science
2026-2027
Note: All curriculum revisions will be updated immediately on the website
http://www.publichealth.gwu.edu

Program Co-Directors

Biostatistics Concentration		**Bioinformatics Concentration**
Guoqing Diao, PhD Professor and Director Department of Biostatistics and Bioinformatics 800 22nd Street, NW, Suite 7560 Washington, DC 20052 Tel: 202-994-2503 Email:gdiao@gwu.edu	Toshi Hamasaki, PhD Professor and Co-Director Department of Biostatistics and Bioinformatics 6110 Executive Blvd, Suite 750 Rockville, MD 20852 Tel: 301-881-9260 Email: thamasaki@gwu.edu	Keith A. Crandall, PhD Professor & Director Department of Biostatistics and Bioinformatics 800 22nd Street, NW, Suite 7670 Washington, DC 20052 Tel: 571-553-0107 Email: kcrandall@gwu.edu

Mission

The PhD in Health Data Science trains the next generation of data science leaders for applications in public health and medicine. Students in the program develop innovative statistical and computational methods for data analysis and for deriving new scientific discoveries in the public health and biomedical sciences. The program takes advantage of the rich biostatistical and bioinformatics resources at GW and in the Nation's Capital and is designed to prepare students to be independent researchers in the development of methodologies, and effective collaborators in interdisciplinary studies.

Competencies

Biostatistics Concentration Competencies

1. **Biostatistical Methodology**: Develop and implement innovative biostatistical methodologies that can be utilized to solve complex problems in and improve efficiency of biomedical and public health research.
2. **Advanced Statistics & Computation**: Implement fundamental and advanced statistical methods and associated computing to effectively and accurately analyze complex public health and medical data.
3. **Communication**: Communicate and collaborate with scientific colleagues such as public health, medical professionals, and research scientists via effective listening, critical questioning, clear written documentation, and oral presentation skills.
4. **Leadership**: Provide biostatistical leadership in the design, conduct, analysis, and reporting of collaborative research studies.
5. **Teaching**: Educate health professionals, research scientists, and students using effective didactic and instructive methodologies to relay complicated mathematical and statistical analyses.

Bioinformatics Concentration Competencies

1. **Computation**: Apply skills in software design and programming to create algorithms for biological data analyses.
2. **Biology**: Apply knowledge of molecular biology, genomics, genetics, evolution, and systems biology through statistics and computation to address novel research questions.
3. **Statistics**: Create modeling and data analyses using statistical and mathematical applications.
4. **Conceptual Integration and Application in Bioinformatics**: Develop project management skills and critical thinking and apply integrated concepts and data across fields of computer science, statistics, data science, biology, and health sciences through bioinformatics.

5. **Independent Research:** Plan and conduct original, independent research through application and integration of skills and knowledge developed through the preliminary phase of the PhD program.

Admission Requirements

Applicants must hold an undergraduate degree from an accredited institution of higher learning exhibiting a strong background in mathematics. Normally, a B average (or equivalent) from an accredited college is required. With evidence of special promise, such as high Graduate Record Examination (GRE) scores, an applicant whose academic record falls short of a B average may be accepted on a conditional basis. **All applicants for the biostatistics track are required to submit current GRE scores (within 5 years of matriculation date)**. GRE scores are optional for the bioinformatics track. Meeting the minimum requirements does not assure acceptance. The department may set higher admission standards and financial support packages available for new students limits the number of applicants accepted. Applicants must provide evidence of the completion of their undergraduate and/or graduate work before registration in the Milken Institute School of Public Health (GWSPH) is permitted. Graduate courses taken prior to admission while in non-degree status may not be transferable into GWSPH programs. The PhD program is designed to serve students with an undergraduate degree and/or a Master's degree. Students completing a master's degree prior to admission to the PhD degree program may transfer up to 24 credits towards the PhD coursework requirements. External credits must have been earned from an accredited institution in the last 6 years with a minimum grade of 3.0 (B) in each course. In this case, 48 hours of additional coursework and dissertation research will be required, including the instructional opportunities.

Concentration-Specific Requirements/Pre-requisites

Biostatistics Concentration	*Bioinformatics Concentration*
• Three semesters of calculus (through multivariable calculus) • A course in linear algebra • A course in undergraduate statistics Additional advanced courses in mathematics and calculus-based probability are encouraged but not a requirement for admission.	• A course in statistics • A course in introductory biology and/or a course in computer programming • Typically, an undergraduate major in either biology, statistics, mathematics, computer science, bioinformatics, and/or bioengineering

All applications are submitted through SOPHAS.org. Information about GWSPH Admissions and policies are available online at http://publichealth.gwu.edu/admissions.

Comprehensive Exam

Biostatistics Concentration:

The Comprehensive Examination is given over a two-week period in the summer before the third (fourth) year of the program for full-time (part-time) students, with a theory examination in the first week and an applications examination in the second week. The theory examination (5 hours, in-class, open-book) covers material in the areas of statistical inference, linear model, and generalized linear model. The applications examination (4 days, take-home, open-book) focuses on applied problems requiring statistical inference based on data analysis. See the HDS PhD (Biostatistics) Comprehensive Exam Guidelines for details. The essential skills for this examination are generally developed through coursework in the first two years of the program. A student who fails to pass the comprehensive examination may, with the approval of the program director, repeat the examination the following year. Failure on the second attempt will result in termination from the PhD program.

Bioinformatics Concentration: The Comprehensive Examination is given by the end of the second semester of the second year of the PhD program (can be earlier for those coming in with a Master's degree and/or equivalent work experience) and consists of a written and oral component.

The written component is an NIH F31 award application focused on the student's current concept of a research project. The written component is developed in coordination with the dissertation advisor.

With advisor approval, the student schedules the oral exam component and delivers the written component to the advisory committee at least one week before the oral examination meeting. The oral exam covers concepts presented in the written component and addresses coursework concepts acquired through the graduate program.

Upon successful completion of the required course work and the PhD Comprehensive Examination the candidate advances to PhD Candidacy: the dissertation research.

Graduation Requirements

All students accepted to this PhD program must follow the policies and regulations outlined in the GW Bulletin and the GWSPH Student Handbook. These include, but are not limited to:

- Minimum Grade Point Average (GPA)
- Minimum Course Grades
- Continuous Enrollment
- Maximum Time to Degree
- Required Minimum Degree Credits
- Transfer Credit Policy
- CITI Training
- Integrity Quiz and Plagiarism
- Professional Enhancement (PE) Hours
- PUBH 6080 Pathways to Public Health
- Graduate Teaching Assistantship Program
- Comprehensive Examination
- Culminating Experience (Dissertation)

Program Requirements

Biostatistics Concentration

The curriculum includes parallel development of theory and applications as well as coverage of specific biostatistical topic areas and ethical issues in the conduct of biostatistical and medical research. The program requires a total of 72 credit hours of coursework and research with a minimum of 12 credits of dissertation research.

Course Distribution Summary	Credits
Required Core Courses	12
Required Concentration Courses	27
Electives NOTE: At least 3 credits must be taken from each of the knowledge areas	12 Minimum
Dissertation Research Note: Research may be taken in 1 credit increments, but a minimum total of 12 credits are required and a maximum of 24 credits can be taken.	12-24
Total credits	**72**

Bioinformatics Concentration

The curriculum includes broad training across core areas of bioinformatics, including statistics, biology, computer science, and ethical issues in the conduct of biomedical research. The program requires a total of 72 credit hours of coursework and research with a minimum of 12 credits of dissertation research.

Course Distribution Summary	Credits
Required Core Courses	12
Required Concentration Courses	18
Electives NOTE: At least 3 credits must be taken from each of the knowledge areas	18 Minimum
Dissertation Research Note: Research may be taken in 1 credit increments, but a minimum total of 12 credits are required with a maximum of 24 credits can be taken.	12-24
Total credits	**72**

Milken Institute School of Public Health THE GEORGE WASHINGTON UNIVERSITY	**Department of Biostatistics and Bioinformatics** **PhD in Health Data Science** **2026-2027** Note: All curriculum revisions will be updated immediately on the website http://www.publichealth.gwu.edu	
Required Core Courses – Both Concentrations (12 credits)		
Course	**Course Name**	**Credits**
PUBH 6080	Pathways to Public Health [1]	0
PUBH 6850	Introduction to SAS for Public Health Research	1
PUBH 6851	Introduction to R for Public Health Research	1
PUBH 6852	Introduction to Python for Public Health Research	1
PUBH 6860	Principles of Bioinformatics	3
PUBH 6886	Statistical and Machine Learning for Public Health Research [2]	3
PUBH 8001	PhD Seminar: Cross-Cutting Concepts in Public Health	1
PUBH 8475	Ethics in Domestic and International Research	1
PUBH 8413	Research Leadership	1
GTAP	GradTeachingAsst Certification (This includes UNIV 0250 Graduate Assistant Certification Course and is required before TA-ing; credits do not count towards PhD degree)	--
Required Courses —Biostatistics Concentration (27 credits)		
PUBH 6864	Applied Survival Analysis for Public Health Research [2]	3
PUBH 6866	Principles of Clinical Trials [2]	3
PUBH 6887	Applied Longitudinal Data Analysis for Public Health Research	3
PUBH 8870	Statistical Inference for Public Health Research I [2]	3
PUBH 8871	Statistical Inference for Public Health Research II [2]	3
PUBH 8875	Linear Models in Biostatistics [2]	3
PUBH 8877	Generalized Linear Models in Biostatistics [2]	3
PUBH 8879	An Introduction to Causal Inference for Public Health Research	3
PUBH 8880	Statistical Computing for Public Health Research	3
Required Courses —Bioinformatics Concentration (18 credits)		
PUBH 6854	Applied Computing in Health Data Science (or equivalent programming course)	3
PUBH 6859	High Performance and Cloud Computing	3
PUBH 6861	Genomics	3
PUBH 6868	Quantitative Methods	3
PUBH 6884	Bioinformatics Algorithms and Data Structures	3
PUBH 8885	Computational Biology	3
Dissertation Research (12 Credits Minimum)		
PUBH 8999	Dissertation Research • Biostatistics concentration • Bioinformatics concentration	 12 – 24 12 – 24

[1] Students without a prior degree from a CEPH-accredited program or school of public health are required to successfully complete the zero-credit, online course Pathways to Public Health (PUBH 6080) within one year of matriculation. There is no fee for this course.

[2] Courses are basis of comprehensive exam for the Biostatistics concentration.

Elective Courses

Biostatistics: 12 Credits Minimum

Bioinformatics: 18 Credits Minimum

Both Concentrations: at least 3 credits in each knowledge area (Biostatistics, Bioinformatics, and Cognate)

Courses required for a concentration may not count towards the elective requirement. Students may add additional elective courses approved in advance by advisor. Electives planned with advisor.

Course	Course Name	Credits
Biostatistics (at least 3 credits)		
PUBH 6262	Introduction to Geographic Information Systems	1
PUBH 6263	Advanced GIS	1
PUBH 6856	Advanced SAS for Public Health Research	1
PUBH 6862	Applied Linear Regression Analysis for Public Health Research	3
PUBH 6863	Applied Meta-analysis	1
PUBH 6865	Applied Categorical Data Analysis	3
PUBH 6888	Introduction to Neural Networks	1
PUBH 8871	Statistical Inference for Public Health Research II	3
PUBH 8877	Generalized Linear Models in Biostatistics	3
PUBH 8878	Statistical Genetics	3
PUBH 8888	Advanced Topics in Clinical Trials	2
STAT 6223	Bayesian Statistics: Theory and Applications	3
STAT 6287	Sample Surveys (Modern Theory of Sample Surveys I)	3
STAT 8257	Probability	3
STAT 8265	Multivariate Analysis	3
STAT 8288	Topics in Sample Surveys (Modern Theory of Sample Surveys II)	3
Bioinformatics (at least 3 credits)		
PUBH 6854	Applied Computing in Health Data Science	3
PUBH 6859	High Performance and Cloud Computing	3
PUBH 6861	Genomics	3
PUBH 6867	Health Data Visualization	3
PUBH 6875	Microbiome Data Analysis in R	3
PUBH 6884	Bioinformatics Algorithms and Data Structures	3
PUBH 6889	Computations for Neural Networks	1
PUBH 6890	Advanced Neural Networks	1
PUBH 6899	Topics in Biostatistics and Bioinformatics (Databases Structure Query Language)	1
PUBH 6899	Topics in Biostatistics and Bioinformatics (Generative Artificial Intelligence)	3
PUBH 8885	Computational Biology	3
PUBH 8886	Population Genetics	3
Cognate (at least 3 credits) *Includes any of the previously listed electives*		
BIOC 6246	Advanced Genomic Data Analysis: From Bulk NGS to Single-Cell	3
CSCI 6221	Advanced Software Paradigms	3
CSCI 6231	Software Engineering	3
DATS 6103	Introduction to Data Mining	3
DATS 6450	Topics in Data Science (Data Learning for Data Science)	3
PSYC 8202	Psychological Research Methods and Procedures	3
PUBH 6003	Principles and Practice of Epidemiology	3
PUBH 6276	Public Health Microbiology	3
PUBH 6278	Public Health Virology	3
PUBH 6894	Research Analytics	3
PUBH 6897	Research in Biostatistics and Bioinformatics	1-4
PUBH 6899	Topics in Biostatistics and Bioinformatics (various)	1-3
PUBH 8244	Doctoral Topics: Cancer Epidemiology	3
PUBH 8250	Doctoral Topics: Epidemiology of HIV/AIDS	3
PUBH 8259	Doctoral Topics: Epidemiologic Surveillance in Public Health	3
PUBH 8419	Measurement in Public Health and Health Services	3
PUBH 8878	Statistical Genetics	3

PUBH 8888	Advanced Topics in Clinical Trials	3
PUBH 8890	Research and Teaching Orientation	1
PUBH 8891	Internship in Data Science	1-3
PUBH 8899	Topics in Biostatistics and Bioinformatics (various)	1-3