1 INTRODUCTION

This handbook describes the Master’s Degree Program in Computational Biology and Quantitative Genetics offered by the Departments of Biostatistics and Epidemiology at Harvard University. The SM program provides students with the rigorous quantitative training and essential skills needed to successfully meet the challenges presented by large-scale public health data – “Big Data” – in biomedical research. The program is designed to prepare students for a career as a bioinformatics analyst or bioinformatics engineer in universities and hospitals, research organizations, and the pharmaceutical and biotechnology industries. It can also provide the foundation for further doctoral studies.

The sections of this handbook include information and regulations concerning entrance requirements, program descriptions, degree requirements, and other program policies. Policies and official requirements of the School of Public Health are set forth in the Harvard T. H. Chan School of Public Health Student Handbook (https://www.hsph.harvard.edu/student-handbook/). Each graduate student is responsible for general knowledge of, and adherence to, the policies and requirements of the degree program in which the student is enrolled. Additional program information is available at the website https://www.hsph.harvard.edu/sm-computational-biology/program/. Vitally important for our community is that all members demonstrate respect for each other and our discipline. For all members of the community, respect is demonstrated by attending all scheduled classes or meetings, and arriving on time, fully prepared, and ready to participate.

This handbook was prepared by the Program Directors and approved by the Executive Committee of the Program in Computational Biology and Quantitative Genetics. The Program Directors are responsible for reviewing the student’s program of study, and have the authority to consider exceptions to the rules and regulations established by the Executive Committee. Recommendations of the Program Directors are forwarded to the Executive Committee for final approval. Both the Program Directors and the Executive Committee welcome suggestions and comments.
PROGRAM LEADERSHIP AND ADMINISTRATION

Executive Committee
Dr. John Quackenbush, Program Co-Director
Email: johnq@hsph.harvard.edu
Dr. Peter Kraft, Program Co-Director
Email: pkraft@hsph.harvard.edu
Dr. Martin Aryee
Dr. Lori Chibnik
Dr. Immaculata De Vivo
Dr. Kimberly Glass
Dr. Adam Haber
Dr. Kyu Ha Lee
Dr. Junwei Lu
Dr. Briana Stephenson

Administration
Jelena Follweiler
Senior Manager of Academic Services, Biostatistics
Phone: 617-432-1087
Email: jtillots@hsph.harvard.edu

David Cruikshank
Graduate Program Coordinator, Biostatistics
Phone: 617-432-5061
Email: dcruihkshank@hsph.harvard.edu

Eric DiGiovanni
Assistant Director of Graduate Studies, Epidemiology
Phone: 617-432-1328
Email: edigiova@hsph.harvard.edu

Career Development
Dr. Erin Lake
Director of Student Development, Biostatistics
Co-Director, Master’s Program in Biostatistics
Email: eklake@hsph.harvard.edu

Websites
Master of Science in Computational Biology and Quantitative Genetics
Department of Biostatistics
Department of Epidemiology
# Master of Science in Computational Biology and Quantitative Genetics

This 80-credit program, offered jointly by the Departments of Biostatistics and Epidemiology, is designed to provide students with:

- The biological background needed to understand and interpret data
- A bioinformatics background providing familiarity with essential tools and data resources
- Computational skills used to analyze and manage “Big Data”
- Statistical skills required to appropriately analyze large quantitative datasets
- Epidemiological skills necessary for the design, conduct, and analysis of experiments

The SM in Computational Biology and Quantitative Genetics is intended as a terminal professional degree which will enable you to launch your career in bioinformatics. It can also provide the foundation for further doctoral studies in biostatistics, epidemiology, computational biology, and other related fields.

Students will receive training in quantitative methods, including linear and logistic regression, survival analysis, longitudinal data analysis, statistical computing, clinical trials, statistical consultation and collaboration, and epidemiology. Students will also gain a strong foundation in modern molecular biology and genetics, computer programming, the use and application of tools for analysis of genomic data, methods for integrative analysis, and meta-analysis of genes and gene function.

## 2.1 Core Competencies

The curriculum for the Master’s in Computational Biology and Quantitative Genetics will provide students with the skills essential to contribute to research projects involving the large, complex genomic datasets that are increasingly common in all areas of biomedical, biological, and public health research. These skills include core competencies in five areas:

1. **Biological Background**
   - Working knowledge of molecular genetics, the structure and organization of the human genome, gene expression regulation, epigenetic regulation, gene functional descriptions, and modern technologies including genotyping, genome-seq, exome-seq, RNA-seq, ChIP-seq, etc, and their applications, and as well understanding of metagenomics.

2. **Bioinformatics Background**
   - Familiarity and ability to use the major genomics data resources, basic knowledge of sequence analysis, familiarity with gene functional annotation and pathway analysis, ability to write data management and analysis scripts, working knowledge of data mining and statistical analysis techniques as well as machine learning approaches, and understanding of modern network modeling techniques.

3. **Computational Skills**
   - Working knowledge of UNIX, a scripting language such as perl or python, an advanced programming language such as c, c++, or java, and R/Bioconductor, and familiarity with database programming and modern web technologies.

4. **Biostatistics Skills**
   - Fundamental understanding of basic statistical inference and applied regression, survival, longitudinal, and Bayesian statistical analysis
5. Epidemiology Skills
   • Ability to critique the existing evidence for a particular research topic, review and summarize information from many studies, to develop robust research questions and design experiments including power calculations, to understand ethical issues in a given study, to assess and correct for measurement errors, and to access and integrate multiple data resources.

2.2 Admissions Procedures and Requirements

2.2.1 Harvard T.H. Chan School of Public Health Requirements

Application for admission to the SM program is available online on the Admissions Office website (https://www.hsph.harvard.edu/admissions/admissions/how-to-apply/application-requirements/). For information on general requirements for admission, contact the Admissions Office by phone (617/432-1031) or through their website (https://www.hsph.harvard.edu/admissions/).

2.2.2 Program Requirements

All candidates for admission to the SM in Computational Biology and Quantitative Genetics program should have successfully completed the following:

   • An undergraduate degree in mathematical sciences or allied fields (e.g., biology, psychology, economics),
   • Calculus through partial differentiation and multivariable integration,
   • One semester of linear algebra or matrix methods,
   • Either a two-semester sequence in probability and statistics or a two-semester sequence in applied statistics,
   • At least one semester of training in biology, with some familiarity with molecular biology and genetics.

In addition, applicants are encouraged to have completed other courses in quantitative areas and in areas of application in the biological sciences. Practical knowledge of computer scripting and programming as well as experience with a statistical computing package such as R is highly desirable. Additional research or work experience is beneficial, but not required. Applicants should show excellence in written and spoken English.

Evidence that these requirements have been fulfilled should form part of the application.

2.3 Intra/Inter-departmental Biostatistics Degree Program Switch Protocol

The Department of Biostatistics, and several other academic departments at the School (such as Epidemiology), offer master’s degree programs. These programs each have their own goals and requirements and make independent decisions about admissions. Students should carefully choose the program to which they apply and we expect students admitted to a program to meet the requirements of that program.

However, we recognize that the interests of some students may change during their time in graduate school. Therefore, the Department has established the following procedure for students applying for a change in program. This applies to students applying for a switch within the Department of Biostatistics, or between the Department of Biostatistics and another department at the School, such as Epidemiology. Please note that transfers between programs are not automatic and may not be approved, and if a student has received a scholarship or other funds from a degree program, that funding will not transfer to the new degree.

1. Students must complete at least one full semester of coursework before applying for a program transfer.
2. Students must enroll in and successfully pass any required coursework for their current program before beginning the transfer process. Please see your current degree program handbook for a list of required coursework. Students may also need to complete coursework in their proposed program so as to not fall behind in requirements to finish their program on time if approved for transfer.

3. Students must complete and submit a formal application and include an updated statement of purpose describing the reason for seeking a program transfer, current CV and one letter of recommendation.

4. The Directors of department master’s programs, including the Directors of the student’s current program and prospective program, will schedule an interview with the students to assess their application. At the meeting, students will provide a copy of their CV and statement of purpose to each Director, and will briefly explain their decision to apply for a transfer. The Directors will consider the applications and render a decision whether to approve or deny the transfer. These meetings will be scheduled in early January for students applying for transfer to start in the Spring semester, and in late May for students applying to transfer after the Spring semester.

5. The Directors will discuss and notify the student of their decision within three days. All decisions are final.

6. If approved, completed paperwork must be submitted to the Registrar’s Office, and then to the Senior Manager of Academic Services.

Note that all program transfers are at the discretion of the Program Directors and are not guaranteed. If approved, students will be assigned a new academic advisor affiliated with the chosen degree program. It is the responsibility of the students to ensure that they complete all requirements for their degree program.

2.4 Curricular Practical Training (CPT) Approval for Students with F-1 Visas

To be considered CPT, the work must not only be related to the student’s major field of study but must also be an integral part of an established curriculum. Before seeking off-campus internship opportunities, students are required to discuss their plans with Elizabeth Capuano (see contact info below) from the Harvard International Office to determine their CPT eligibility. Please note that CPT eligibility may be impacted by the March 2020 guidance issued by the Department of Homeland Security - Student Exchange Visitor Program.

There are two ways in which students are eligible for CPT:

1. Employment that is a required part of a degree program, such as a required internship or practicum. This requirement must be formally documented in school publications, such as a student handbook.

2. Employment that is not required by a degree program, but for which a program will award academic credits. This could include training courses such as a field studies course, an independent study course that is based on an internship.

Our program does not offer the first option above. However, we do allow the second option with approval from either of the program directors. International students who wish to pursue this option MUST speak with Elizabeth Capuano (elizabeth.capuano@harvard.edu), our representative at the Harvard International Office (HIO), before beginning interview processes at prospective internship sites to discuss the requirements for CPT authorization. Students should also speak to one of the directors about whether their employment would qualify for academic credits, as the academic credits are required for CPT authorization. If CPT eligibility is established, please note that students MUST obtain CPT authorization PRIOR to beginning the internship.

The most up-to-date information about CPT will be found here: http://www.hio.harvard.edu/curricular-practical-training-cpt.
2.5 Advising and Degree Program Approval

2.5.1 Academic Advisor

All entering students are assigned an academic advisor to help plan course loads and explain program requirements. At the earliest possible date, the student and the academic advisor will develop a program of study. Students should bring their Master's Degree Program form (Section 2.5.2) to all meetings with their advisors to keep on track with their requirements. Should a student wish to change his/her academic advisor, he or she is encouraged to discuss this with the Program Directors. In addition, SPH provides services for all students with clinically documented learning and/or physical disabilities.

2.5.2 Departmental Approval of Program

The Master’s Degree program plan must be submitted to the student’s academic advisor and the Program Directors for approval, using the Master’s Degree Program form provided at least one semester prior to their expected graduation date (see timeline in Section 3 for details).

2.5.3 Epidemiology Requirement

The School of Public Health requires that Master’s students must successfully pass one epidemiology course. The program requires that EPI 201 be taken to satisfy this requirement.

2.5.4 Public Health Practice Requirement

Students may be required to take a public health course by the School of Public Health as part of their accreditation requirements. These requirements will be communicated to all incoming students by the School of Public Health directly.

2.5.5 Research Ethics Requirement

Students must satisfy a research ethics requirement by completing a course in responsible conduct of research or by completing an online training course during the first year in the program (see the timeline in Section 3 for details). Students who feel they have already completed an equivalent training program must submit adequate documentation to, and receive approval from, the Senior Manager of Academic Services in Biostatistics (see contact list on page 1) during the first semester in residence.

2.6 Satisfactory Progress Requirements

For students in the SM2 program, a total of 80 credits are required with a minimum of 55 ordinal credits from the core courses, tracks, and electives listed in Section 2.7.1. In addition, SPH students must remain in good academic standing, must complete program requirements within the designated time to degree, and must maintain a cumulative average of 2.70 or above. All ordinal grades for courses used to satisfy program requirements specified in Section 2.7.1 must be at the level of B- or higher. Courses taken on a pass/fail basis cannot be used to satisfy ordinarily graded program requirements.

A detailed presentation of SPH’s regulations for Master’s students is found at https://www.hsph.harvard.edu/student-handbook/. All Master’s students and their advisors should make sure that SPH and CBQG program requirements are met according to schedule.

2.7 Degree Requirements

A total of 80 credits are required for the SM2 in Computational Biology and Quantitative Genetics. A minimum of 55 ordinal credits of coursework must be taken from the core courses, tracks, and electives listed below. Students with prior equivalent background to any of the required courses or strong reasons to take a different course can request permission from the Program Directors for a substitution of one or more of the required courses. Students wishing to substitute courses from other departments or institutions

6
for those listed here must obtain prior approval from their academic advisors and one of the program co-directors (see more detailed note after course listings). Although academic advisors and departmental staff will work with students to monitor progress, it is ultimately each student’s responsibility to ensure that all requirements are met.

2.7.1 Course Requirements for the SM2 Degree

Fifty-five credits of ordinally graded courses must be taken from the following courses. This includes a 12.5 credit ordinally graded core curriculum consisting of:

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credits</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>BST 210</td>
<td>Applied Regression Analysis</td>
<td>5</td>
<td>Fall</td>
</tr>
<tr>
<td>BST 280</td>
<td>Introductory Genomics &amp; Bioinformatics for Health Research</td>
<td>2.5</td>
<td>Fall 2</td>
</tr>
<tr>
<td>EPI 201</td>
<td>Introduction to Epidemiology Methods I</td>
<td>2.5</td>
<td>Fall 1</td>
</tr>
<tr>
<td>EPI 249</td>
<td>Molecular Biology for Epidemiologists</td>
<td>2.5</td>
<td>Fall 1</td>
</tr>
</tbody>
</table>

An additional ten credits comprised of courses in either one of the two following tracks:

**Statistical Genetics Track**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credits</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>BST 227</td>
<td>Introduction to Statistical Genetics</td>
<td>2.5</td>
<td>Fall 2</td>
</tr>
<tr>
<td>EPI 293</td>
<td>Analysis of Genetic Association Studies</td>
<td>2.5</td>
<td>Wintersession</td>
</tr>
<tr>
<td>EPI 507</td>
<td>Genetic Epidemiology</td>
<td>2.5</td>
<td>Fall 2</td>
</tr>
<tr>
<td>EPI 511</td>
<td>Advanced Population and Medical Genetics</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>or</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 535</td>
<td>Epidemiologic Challenges to the Interpretation of Genetic Analyses</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

or

**Computational Biology Track**

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credits</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>BST 281</td>
<td>Genomic Data Manipulation</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 282</td>
<td>Introduction to Computational Biology and Bioinformatics</td>
<td>5</td>
<td>Spring</td>
</tr>
</tbody>
</table>

A minimum of 22.5 additional credits will come from the field of study tracks above or the following list of elective courses:

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credits</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>BST 212</td>
<td>Survey Research Methods in Community Health</td>
<td>2.5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 214</td>
<td>Principles of Clinical Trials</td>
<td>2.5</td>
<td>Spring 1</td>
</tr>
<tr>
<td>BST 217</td>
<td>Statistical &amp; Quantitative Methods for Pharmaceutical Regulatory Services</td>
<td>2.5</td>
<td>Spring 2</td>
</tr>
<tr>
<td>BST 221</td>
<td>Applied Data Structures and Algorithms</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>BST 222</td>
<td>Basics of Statistical Inference</td>
<td>5</td>
<td>Fall 1</td>
</tr>
<tr>
<td>BST 223</td>
<td>Applied Survival Analysis</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 226</td>
<td>Applied Longitudinal Analysis</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 228</td>
<td>Applied Bayesian Analysis</td>
<td>5</td>
<td>Fall</td>
</tr>
<tr>
<td>BST 230</td>
<td>Probability Theory and Applications I</td>
<td>5</td>
<td>Fall</td>
</tr>
<tr>
<td>BST 231</td>
<td>Statistical Inference I</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 232</td>
<td>Methods I</td>
<td>5</td>
<td>Fall</td>
</tr>
<tr>
<td>BST 260</td>
<td>Introduction to Data Science</td>
<td>5</td>
<td>Fall</td>
</tr>
<tr>
<td>or</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS 109A</td>
<td>Data Science 1: Introduction to Data Science</td>
<td>5</td>
<td>Fall, aka STAT 121A</td>
</tr>
<tr>
<td>BST 261</td>
<td>Data Science II</td>
<td>2.5</td>
<td>Spring 2</td>
</tr>
<tr>
<td>BST 262</td>
<td>Computing for Big Data</td>
<td>2.5</td>
<td>Fall 2</td>
</tr>
<tr>
<td>BST 263</td>
<td>Statistical Learning</td>
<td>5</td>
<td>Spring</td>
</tr>
<tr>
<td>BST 267</td>
<td>Introduction to Social and Biological Networks</td>
<td>2.5</td>
<td>Fall 2</td>
</tr>
<tr>
<td>Course Code</td>
<td>Course Title</td>
<td>Credits</td>
<td>Term</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>BST 270</td>
<td>Reproducible Data Science (2.5 credits, Wintersession)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BST 273</td>
<td>Introduction to Programming (2.5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BST 283</td>
<td>Cancer Genome Data Science (5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 202</td>
<td>Elements of Epidemiologic Research: Methods 2 (2.5 credits, Fall 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 203</td>
<td>Study Design in Epidemiologic Research (2.5 credits, Spring 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 204</td>
<td>Analysis of Case-Control and Cohort Studies (2.5 credits, Spring 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 221</td>
<td>Pharmacoepidemiology (2.5 credits, Fall 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 271</td>
<td>Propensity Score Analysis (1.25 credits, Wintersession)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 288</td>
<td>Data Mining and Prediction (2.5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EPI 289</td>
<td>Models for Causal Inference (2.5 credits, Spring 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID 271</td>
<td>Advanced Regression for Environmental Epidemiology (2.5 credits, Spring 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDS 280</td>
<td>Decision Analysis for Health and Medical Practices (2.5 credits, Fall 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDS 282</td>
<td>Economic Evaluation of Health Policy &amp; Program Management (2.5 credits, Spring 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RDS 285</td>
<td>Decision Analysis Methods in Public Health and Medicine (2.5 credits, Spring 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOPHYS 170</td>
<td>Quantitative Genomics (5 credits, Fall, aka MIT HST .508)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOPHYS 376</td>
<td>Functional &amp; Computational Genomics Studies of Transcription Factors &amp; Cis Regulatory Elements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 701</td>
<td>Foundations in Biomedical Informatics I (5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 702</td>
<td>Foundations of Biomedical Informatics II (5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 703</td>
<td>Precision Medicine I: Genomic Medicine (2.5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 705</td>
<td>Precision Medicine II: Integrating Clinical and Genomic Data (2.5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 706</td>
<td>Data Visualization for Biomedical Applications (2.5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 713</td>
<td>Computational Statistics for Biomedical Sciences (2.5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOMI 715</td>
<td>Computing Skills for Biomedical Informatics (2.5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMIF 201</td>
<td>Concepts in Genome Analysis (5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS 181</td>
<td>Machine Learning (5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>or</td>
<td>MIT 6.867 Machine Learning (5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT 6.862</td>
<td>Applied Machine Learning (5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT HST .956</td>
<td>Machine Learning for Healthcare (5 credits, Spring)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIT 6.878</td>
<td>Computational Biology: Genomes, Networks, Evolution (5 credits, Fall)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1 May be taken concurrently with BST 210 in the fall if in the BIO or CBQG degree programs. This course is a prerequisite for the doctoral-level core Biostatistics theory courses.

2 These courses are strongly recommended.

Other advanced courses in Biostatistics, including many of the special topics or Wintersession courses, and courses at MIT and the Faculty of Arts and Sciences (FAS) that are offered at an advanced level, may also be acceptable. Students wishing to substitute courses from other departments or institutions for those listed here must obtain prior approval from their academic advisors and one of the program co-directors. Courses not receiving prior approval will not count toward meeting the program requirements. No more than three courses outside the list above will be approved without special consideration. If you need cross-registration credit conversion, see: https://www.hsph.harvard.edu/r-o-student-knowledge-center/. To request a substitution, students should email the Program Directors with their reasons for the request (please cc the graduate program coordinator, David Cruikshank, so that he may file the information for the final degree audit).

2.7.2 Collaborative Research Thesis

The Collaborative Research Thesis provides students with valuable real-world experience doing research and working with data in Boston’s premier biomedical institutions. In the thesis, a student will perform activities related to the design, conduct, and analysis of research studies with the goal of addressing a relevant question requiring the use of methods in computational biology or quantitative genetics. For the thesis, students are
generally mentored by a member of the program faculty or other affiliated quantitative scientist with a faculty-level appointment at SPH. Students can also conduct research mentored by members at Harvard University, or a Harvard-affiliated hospital, or other local organizations including nonprofit organizations, biotech startup companies, or pharmaceutical companies, provided they also have a co-mentor among the program faculty (Drs. Kraft or Quackenbush are generally willing to provide such mentorship). Students are responsible for finding and selecting their own advisors and working to develop a suitable research program. A link to a list of possible thesis advisors is provided here, although this list should not prevent you from exploring other potential advisors. Students may also ask their academic advisor or the Program Directors about possibilities for thesis research and advisors; Dr. Erin Lake also has many industry contacts, including those looking for possible SM candidate interns.

Students should notify the Graduate Program Coordinator and the Senior Manager of Academic Services in Biostatistics (David and Jelena; see contact list on page 1) by e-mail about their thesis advisor choice by the end of their first academic year. Therefore, students should research and speak to potential advisors and thesis topics throughout their first year. Please cc the thesis advisor in the e-mail so there is a record that they have agreed to serve and please note if there is a co-advisor from the program.

Before the start of the Fall semester of their second year, students should work with their thesis advisors to prepare a short thesis proposal outlining their research question, the data they will use, the methods that will be applied, and the anticipated results; any questions regarding human subjects research should also be addressed. The thesis proposal should have section headings reflecting these four (or five) required elements as well as references cited and should be about two pages in length (although slightly longer is fine) but it should provide sufficient detail to allow the program directors to assess the proposed project. Once the proposal has been approved by the Program Directors, send the approved proposal to David and Jelena. Examples of previous successful proposals may be obtained from David and/or Jelena to aid in proposal development.

As part of this thesis requirement, students must register for a 10-20 credit ordinarily graded CBQG Collaborative Research Thesis (CBQG 325) section, normally undertaken during the third and/or fourth semester, after the required core course work has been completed and should generally correspond to the time during which research is conducted, however students may begin their thesis research during the summer following their second semester (see timeline in Section 3 for details).

The student will then write a Master’s thesis of approximately 15-25 double-spaced pages (excluding tables, figures, and references) that describes, in a standard scientific writing style, the medical or public health problem of interest as well as the analytical methods used and their appropriateness, summarizes the data analyses, and provides a scientific interpretation of the data. The student will also orally present this work in a presentation of approximately 30 minutes in length. The defense must be booked for forty-five minutes to allow for questions and discussion by committee after the presentation.

The Master’s thesis and oral presentation will primarily be the work of the student, with only advisory input from the thesis committee members. The Master’s thesis and oral presentation will be evaluated by a thesis committee consisting of a minimum of three members. The members will include the student’s thesis advisor(s), one of the Program Directors, and other Biostatistics faculty members or surrogates as needed (potentially including the student’s academic advisor). The student will submit a Thesis Committee Nomination form before scheduling the thesis defense.

If the thesis work has already been accepted as a journal article, a student may still use this as part of the thesis, but should write introduction and conclusion sections placing the work in context.

The thesis defense should be scheduled by contacting David (and/or Jelena) by late March or early April in their final semester (see timeline in Section 3 for more details). Once a date has been chosen, the Thesis Scheduling form should be completed and submitted to David, and the draft of the Master’s thesis must be submitted to the thesis committee at least two weeks prior to the scheduled oral presentation. After the successful defense, and any moderations requested by the Thesis Committee, the final thesis should be
submitted to the Program Directors, the Thesis Committee, and David (and/or Jelena) by June 1.
3 ADMINISTRATIVE TIMELINE

Detailed requirements and deadlines for degree completion are given on the Harvard T.H. Chan School of Public Health webpage. All forms linked below are also located on the last page of this Graduate Student Handbook.

- **Summer Before Entering Program**
  - Take online course in R. One option is [https://www.datacamp.com/courses/free-introduction-to-r](https://www.datacamp.com/courses/free-introduction-to-r)

- **Year One**
  - **First Semester**
    - Complete Research Ethics requirement by taking the free online CITI program or attending HPM 548 (See Section 2.5.5). You will receive a reminder about where to find the online course before beginning the Fall term.
    - If planning to waive courses, ask for those waivers of fall core courses (BST 210 or BST 280) or the track courses by emailing the Program Directors and David (and/or Jelena) in Biostatistics with details about the course(s) taken or experience that you have that may qualify you for a waiver. To waive EPI 201 which is a school-wide core requirement, you must submit this Waiver of Core Courses Form to the instructor teaching the course for approval, and then to the Registrar’s Office. You may not waive out of EPI 249 without a strong background in the material covered.
  - **Second Semester**
    - Complete or waive spring track courses.
    - Attend scheduled meeting about thesis advisors and research projects (March or April).
    - Search for thesis advisor (see Section 2.7.2) and potential project.
    - Notify David (and/or Jelena) of your thesis advisor choice by May 15.
  - **Summer**
    - May start thesis research over the summer. Signing up for thesis credit (CBQG 325) is not necessary or advisable during the summer.
    - Work on your thesis proposal with your thesis advisor and submit the proposal by form in email to the Program Directors by the beginning of Fall term. Once the proposal has been approved by the Program Directors, send the approved proposal to David (or you may cc David and/or Jelena on the original email to the directors). Examples of previous successful proposals may be obtained from David or Jelena prior to submission.

- **Year Two**
  - **Third Semester**
    - Continue to complete any necessary coursework.
    - Start or continue thesis research and sign up for research credit (CBQG 325) at this time, counting the summer work as a part of this enrollment. The grade for CBQG 325 in the fall may be “incomplete” until you finish your thesis and defend it in the spring term.
Fourth Semester

- Continue to complete any necessary coursework.
- Start or continue thesis research and sign up for research credit (CBQG 325).
- Turn in your final program form by February 15.
- Choose your thesis committee members, and complete this form. Submit it to David (and/or Jelena) by the end of Spring 1 term (no later than March 15).
- Work with David (and/or Jelena) to schedule a room for your thesis defense date/time. The Program Directors should have dates/times blocked off in late April or early May for all thesis defenses. Fill out the Thesis Scheduling form after locking down a date.
- Submit final version of the defended thesis to David (and/or Jelena) by June 1 and let them know if you are willing to share your final version with future students.

4 PROGRAM FORMS

- **CBQG SM2 Degree Program Form**
  https://content.sph.harvard.edu/biostats/publications/cbqg_handbook/SM2_Degree_Form_CBQG.pdf

- **CBQG Thesis Proposal Form**

- **Thesis Committee Nomination Form**
  https://content.sph.harvard.edu/biostats/publications/cbqg_handbook/Thesis_Co