Department of Biostatistics and Bioinformatics

Proposal for Graduate Program in Biostatistics

November 2013
# Department of Biostatistics and Bioinformatics
## Proposal for Graduate Program in Biostatistics

November 2013

## Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>BB PhD Proposal</td>
<td>2</td>
</tr>
<tr>
<td>Appendix A Letters of Support</td>
<td>44</td>
</tr>
<tr>
<td>Appendix B Duke Masters in Biostatistics Curriculum</td>
<td>61</td>
</tr>
<tr>
<td>Appendix C Formal Review of Masters of Biostatistics Program</td>
<td>70</td>
</tr>
<tr>
<td>Appendix D Comparative Programs</td>
<td>80</td>
</tr>
<tr>
<td>Appendix E Budget Summary</td>
<td>82</td>
</tr>
<tr>
<td>Appendix F Graduate Faculty</td>
<td>83</td>
</tr>
<tr>
<td>Appendix G Timeline for Student Progression</td>
<td>86</td>
</tr>
<tr>
<td>Appendix H Curriculum Vitae of Faculty Who Will Participate In the Program</td>
<td>90</td>
</tr>
<tr>
<td>Appendix I Letters of Support from Nancy Andrews, Dean of the School of Medicine</td>
<td>162</td>
</tr>
<tr>
<td>Appendix J Student Learning Assessment Process Template</td>
<td>164</td>
</tr>
<tr>
<td>Appendix K Grants on which B&amp;B Faculty are Principal Investigators</td>
<td>168</td>
</tr>
</tbody>
</table>
Department of Biostatistics & Bioinformatics
Proposal for Graduate Program in Biostatistics

Mission: To educate tomorrow’s leaders in the development and application of statistical methods of biomedical research for the continual improvement of human health.

Executive Summary
The Department of Biostatistics and Bioinformatics is proposing a new PhD program to respond to the current shortage of biostatisticians needed for medical research in academia, government, and industry. To be successful, these researchers need a diverse skill set. First, they need a firm grasp of biostatistical theory and techniques, including a deep knowledge of probability and statistical inference, in order to develop innovative and methodologically sound solutions to complex problems arising in biomedical research. Second, a core characteristic of biostatistical research is that it is responsive to inferential problems in medical research. Thus, an effective biostatistician will have a good grasp of basic biomedical systems and will have the background to understand and translate biomedical problems into a quantitative framework that is amenable to statistical inference. Finally, the need for effective communication is pervasive across all biostatistical research, from problem formulation, to articulating a solution, disseminating results, and educating the next generation of biostatistical researchers. Our core curriculum ensures that each student will develop these fundamental competencies.

The following points highlight the features of our program. We begin with the curriculum and for clarity we divide the points into (1) features of the curriculum that are often found in top biostatistics programs, and (2) novel features of the curriculum that extend the training of our students beyond that found in a traditional biostatistics program. We then list other important program features.

Curriculum:
(1) Features of curriculum that are consistent with training found in traditional biostatistics programs:
- Program is comprised of a structured two-year core curriculum, followed by elective courses and a dissertation.
- The core contains a first year, master’s level sequence in probability and statistical inference, followed by a second year sequence in measure theoretic probability and advance statistical inference, currently being offered in the Department of Statistical Science (DSS). Please see Appendix A containing a letter of support from Alan Gelfand, Chair of DSS, assuring enrollment in these courses for our students.
Similar to other PhD programs in Biostatistics, the core contains specialized training in categorical data analysis, survival analysis, generalized linear models, analysis of correlated and longitudinal data, and theory of linear models.

Throughout the core curriculum, principles of epidemiologic studies are covered in detail.

The core contains a course emphasizing the ethical issues in the conduct of statistical and medical research.

Coursework is required in a ‘cognate’ field, a substantive field outside of biostatistics or statistics. Examples include: epidemiology; biology; biophysics; environmental health; genetics; etc. The cognate field should be complementary to the student’s biostatistical area of interest. Appendix A provides letters of support from several supporting programs.

(2) Novel features of the curriculum that extend the training of our students beyond that found in a traditional biostatistics program:

- The core contains a unique first year ‘problems in biostatistics’ survey seminar. The seminar will involve a gradual progression from required attendance in the first year to active collaborative involvement in the second year, to collaborative presentations in the third year, to independent presentations in fourth and fifth years.

- The core contains a novel course sequence emphasizing biomedical concepts and communication.

- Core contains a novel statistical methods for learning and discovery course which introduces bioinformatic techniques.

- Core contains a novel survey course of modern inferential techniques and theory (to be developed) targeted specifically to advanced graduate students.

- Embedded throughout the curriculum are examples of conflict of interest situations faced by biostatisticians, along with principles of reproducible research and strategies for implementation.

Other academic features of the proposed program:

- Research-oriented experiences, whereby students become affiliated with a faculty member for a period of two to four months to learn about research opportunities and also the practical aspects of a career in biostatistics, are required following the first year of coursework.
• Students will be assigned an initial advisor who will guide the student through selection of faculty affiliations and will mentor the student regarding other issues related to the program.
• Students will engage in the rich academic atmosphere of the department and will be required to attend departmental seminars and to actively participate in departmental journal clubs (clinical trials, statistical genetics, bioinformatics, causal inference).
• Students must pass a comprehensive exam covering the material in the first year core sequence. In order for advanced students to test out of specific first year courses, they must pass the material covered in our first year of theory on the comprehensive exam.
• Students must pass a PhD qualifying exam covering all of the material in the core sequence (typically at the end of their second year).
• In year 3 students take additional electives and select a dissertation advisor, topic, and committee. The committee will consist of the dissertation advisor, at least two faculty members from B&B or DSS, a faculty member from the interdisciplinary cognate program, and, when appropriate, a member of the biomedical area from which the dissertation was derived. Thesis committees will consist of at least four members of the Graduate Faculty in accordance with Graduate School policy.
• Students must typically pass their preliminary exam (dissertation proposal defense) by the end of the third year.
• Dissertation research/writing typically occurs in years 3-5, followed by defense.

**Finances**

We anticipate enrolling 20-21 masters students and 4 PhD students each year. A brief, very conservative, synopsis of the budgeted financial situation at steady state, when fully five cohorts of PhD students have been enrolled, is as follows:

- Tuition revenue from 42 students: approximately $1.5 Million
- Expenses – approximately $1.5 Million
  - Faculty Teaching (including advising): $443,000
  - PhD tuition, fees and stipends: $432,000
  - Director of Graduate Studies: $58,000
  - Masters Scholarships: $150,000
  - Administration: $71,000
  - Other operating: $159,000 (including retreat and G&A)
  - Space: $103,000

This budget anticipates funding four to five PhD students in their second year, with a budgeted contingency of $200,000 per year for backstopping students in year 3 and above, while maintaining a program reserve of $1.1 Million.
I. Rationale for the new program

I.1. Introduction

The Department of Biostatistics and Bioinformatics is proposing a PhD program in Biostatistics. The program will train students to (a) develop innovative new statistical methodology to meet the increasingly complex needs of current medical research; (b) apply existing statistical methodology to biomedical research in a thoughtful and creative manner; and (c) be effective collaborators in today’s multidisciplinary biomedical research environment.

The new PhD program responds to the current shortage of biostatisticians who are prepared to meet the demand for multidisciplinary collaboration in academia, government, and industry. Further, the “big data” explosion in genomics and medical informatics (electronic health records, EHR) will increase demand, particularly for biostatisticians who can pioneer new methods research. In addition, Congress recently established the Patient-Centered Outcomes Research Institute for the purposes of extending our understanding and quantifying the sources of variability in patient-reported outcomes, which will add yet another dimension to the EHR. In examining the potential value of “big data” to the world economy, economic analysts contend that a shortage of talent with statistical expertise and machine learning will significantly constrain the ability of government, business and industry to realize its value: (http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation).

Additionally, the National Institute of General Medical Sciences, which supports most NIH-funded PhD training, has recently instituted a requirement that students have quantitative training. The most broadly applicable training is in biostatistics, which crosses all disciplinary lines. Through our progressive seminar and our core course in communication, our proposed program will train students who then can teach other biomedical trainees.

Demand is strongest for doctoral-level biostatisticians with a relatively unusual constellation of skills. The first of these includes deep analytical competency, which at the doctoral level includes not just the ability to apply existing methodology but also the ability to develop new methods. The second element of this constellation includes fundamental biomedical knowledge, which is necessary for appropriately translating a research question into a statistical one. The third element is communication, problem-framing and problem-solving.
These skills allow the biostatistician to frame a statistical solution in a fashion that is biomedically sound and mathematically tractable.

Our program will be unique in addressing all three areas described above. Moreover, it builds upon an innovative Masters’ program that provides much of the proof of concept for the PhD program. Our vision is that Duke will become recognized as a national leader in training outstanding biomedical statisticians – individuals who combine methodological sophistication with firm grounding in the biomedical disease process, who also have honed leadership and communication skills in order to maximize the impact of their efforts -- toward the ultimate goal of improving public health.

I.1.1. Description of the Department of Biostatistics and Bioinformatics

The Department of Biostatistics and Bioinformatics (B&B) was established in 2000, with the expectation of becoming a national leader in these scientific disciplines. A founding document for the department states: “As the practice of medicine becomes increasingly evidence-based and the wealth of information within medical and genomic databases grows dramatically, the core disciplines of biostatistics and bioinformatics are becoming increasingly important in medical research and training. Scientists ranging from molecular biologists studying complex diseases such as cancer, diabetes, and dementia to clinicians comparing the effectiveness of treatment regimens must build their research protocols on the concepts and methods of these quantitative disciplines. Duke University Medical Center (DUMC) recognized the importance of an academic home to support these disciplines by creating a Department of Biostatistics and Bioinformatics (B&B) and by making a commitment to become a national leader in areas in which it already had numerous strengths but previously lacked adequate infrastructure.”

B&B currently has approximately 50 full-time PhD-trained faculty members, a significant number of whom teach in the two masters-level programs administered by B&B: the Clinical Research Training Program (CRTP), which awards the Master of Health Sciences in Clinical Research, and the Master of Biostatistics program. The current proposal represents a departmental and School of Medicine priority and builds on our successful experience in developing and administering these programs.

I.2. Description of the Field

I.2.1. Working definition of biostatistics
Biostatistics is the application of statistical principles to biomedical questions. In general, biostatistics uses much the same core sets of probability and inference concepts as does applied
statistics. These concepts include distributions and variation, statistical modeling approaches for analyzing data in the presence of variation, and understanding and testing underlying assumptions. The subject-matter knowledge that the biostatistician must learn in order to be successful is biomedical.

Biostatistics underlies modern medical research in each step of scientific inquiry from the research bench to the hospital bedside to the community. Biostatistics is concerned with the development and proper application of methods for study design, data measurement, data generation, and data analysis. Biostatistical methods are used to help understand data by quantifying variation and by separating signal from noise. The fundamental elements of data and variation are ubiquitous in medical research, being found in such areas as cell regulation, gene expression, genetic susceptibility, pharmacokinetics, response to therapy, assessment of medical treatments and new technology, adherence to guidelines, and program evaluation.

I.2.2. Role of biostatistics

Biostatisticians design studies, oversee their implementation, analyze study data and interpret the results. Importantly, they develop new methodologies in these areas. With grant funding becoming more competitive and grant study sections demanding more statistical rigor, those biostatisticians who perform high-profile methodological research lend name recognition to grant applications on which they are co-investigators. Demand is particularly high for biostatisticians who can become facile with the underlying biomedical content, develop innovative solutions that reflect the unique nature of specific problems, and communicate the results to a variety of audiences, including students. Biostatisticians are needed in academia, industry and government.

I.2.3. Applications of biostatistics

As an illustration of the wide application of biostatistics, some examples are provided. First, in the area of clinical trials, ethical imperatives require extracting as much information from as few patients as possible, placing no more patients than necessary into the group with the least effective intervention, and drawing conclusions as rapidly as possible so that they can then be applied to other patients with the same conditions. These ethical imperatives have led to the development of adaptive designs, and the improvement of such designs is a topical area of inquiry. The statistical issues in question become increasingly complex when genomic information is added.

A second example comes from observational registries of patients who have been diagnosed with certain diseases. Such registries -- now feasible with increases in computing power and
the ability to link datasets – can potentially be used to identify rare adverse events associated with certain therapies and to screen for unexpected drug interactions. The risk of false positive conclusions increases with the number of queries that are performed. Drawing appropriate conclusions in the presence of potential false positive results is an area of ongoing statistical inquiry and is an area where insights from genomic statistics are applicable. With the proliferation of electronic health records, such investigations will expand and provide the infrastructure for the design and analysis of more extensive uses of health data. By developing statistical methodology for mining such databases, insights can be generated to support personalized medicine.

Another use of registries and electronic health records is in comparative effectiveness research, which attempts to determine which interventions are most effective in actual practice. Randomization is absent, and conclusions must take into account the fact that groups can differ not only in characteristics that can be measured, but also in characteristics that are unmeasured. This research requires sophisticated understanding of the developing field of causal inference.

Other types of big data that require new or improved statistical approaches include genetic and genomic data as well as imaging data. The availability of millions of data points on relatively small numbers of patients represents a challenge to classical statistical inference.

These examples illustrate that the demand for biostatisticians is likely to grow; that advancing the field requires fundamental understanding of the biomedical problem being studied and that collaboration and communication skills will be essential.

I.2.4. Benefits to Duke of a PhD program in biostatistics

First, a PhD program would benefit efforts in recruiting and retaining top-level talent to join our biostatistics faculty, by providing the opportunity to teach high-level courses, to mentor and publish with excellent graduate students, and to collaborate with an increasing pool of biostatisticians on methods development. Moreover, the program would remove a disincentive that currently exists when we compete for academic biostatisticians with other institutions that have doctoral programs.

A PhD program would also strengthen the existing Masters in Biostatistics program by enhancing the academic atmosphere. Upper level students are a valuable resource for answering questions and providing advice. They also serve as role models and motivation. A by-product of this enhanced learning environment will be a more valuable masters degree and hence increased name recognition for both B&B and Duke in general.
The demand for PhD-level biostatisticians is particularly acute in industry, especially the pharmaceutical industry, and our program would provide highly trained biostatisticians in positions where critical shortages exist. Our current Masters in Biostatistics program includes on its Professional Advisory Council several biostatisticians from pharmaceutical companies and clinical research organizations. Their guidance has helped us tailor our program to their needs. Further, we are currently working with several pharmaceutical companies who are considering providing scholarships in their names, but without obligations.

An additional benefit to Duke relates to the synergy of this program with the Computational Biology and Bioinformatics (CBB) program and the PhD program in the Department of Statistical Science (DSS). Students enrolled in CBB or DSS will benefit from having access to Biostatistics courses that are potentially relevant to their fields of study.

**I.3. Resources at Duke for an innovative program in biostatistics**

The Biostatistics and Bioinformatics Department at Duke has created two successful masters programs that not only serve as proof of concept, but also provide the infrastructure for a more advanced degree. We have demonstrated that B&B has a number of faculty members who are willing and capable to teach classes, support and mentor students, and otherwise contribute to the program. Most importantly, Duke has a world-class medical research enterprise, involving close collaboration between biostatisticians and other researchers, and thus providing a wide variety of projects and experiences that can serve as the program’s “laboratory”.

**1.3.1. Master of Health Sciences in Clinical Research**

Teaching programs in the Department originated with the Clinical Research Training Program (CRTP), which awards a Master of Health Sciences in Clinical Research degree through the School of Medicine. Its target audience is clinical house staff, faculty, and medical students. One of the first programs of its kind in the country, CRTP has served as a model for similar programs at other institutions and now includes a distance learning component; Clinical fellows and other health professionals at the National Institutes of Health (NIH) in Bethesda, Maryland and the Brazilian Clinical Research Institute (BCRI) in Sao Paulo, Brazil participate in the degree program, taking courses remotely in real time via interactive videoconferencing. Since the program’s inception in 1986 as the Biometry Training Program, over 1,150 individual trainees have enrolled in program courses. As of November of 2012, degrees have been awarded to 351 trainees.
Of particular relevance to this document is the number of B&B faculty members that teach in CTRP, 15 during the last two academic years alone. In addition, 26 members of the faculty have served on Examining Committees for degree candidates. The Master of Biostatistics program and our proposed PhD program obtain significant synergies by leveraging this educational infrastructure and experience.

I.3.2. Masters in Biostatistics

Our Masters in Biostatistics program, approved as a professional degree through the Duke School of Medicine, accepted its first class of students in the Fall of 2011. Requirements for admission include good performance in multi-variable calculus, with additional training in areas such as linear algebra and real analysis preferred. Apart from quality of mathematical training, the Admissions Committee also takes into account the school at which training was obtained, a personal statement, and three letters of recommendation. Comparing our 2011-2012 and 2012-2013 students with other masters programs at Duke demonstrates that our new program compares well with these other programs; we are also tightening our requirements.

<table>
<thead>
<tr>
<th></th>
<th>Biostatistics</th>
<th>Biomedical Engineering</th>
<th>Economics</th>
<th>Global Health</th>
<th>Medical Physics</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011-2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRE Verbal</td>
<td>528 [64%]</td>
<td>596 [82%]</td>
<td>569 [77%]</td>
<td>553 [70%]</td>
<td>490 [53%]</td>
</tr>
<tr>
<td>GRE Quantitative</td>
<td>764 [82%]</td>
<td>772 [84%]</td>
<td>776 [86%]</td>
<td>692 [60%]</td>
<td>740 [74%]</td>
</tr>
<tr>
<td>UGPA</td>
<td>3.4</td>
<td>3.7</td>
<td>3.6</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td>2012-2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRE Verbal</td>
<td>593 [78%]</td>
<td>159 [80%]</td>
<td>160 [83%]</td>
<td>158 [77%]</td>
<td>153 [57%]</td>
</tr>
<tr>
<td>GRE Quantitative</td>
<td>776 [85%]</td>
<td>162 [85%]</td>
<td>164 [90%]</td>
<td>157 [71%]</td>
<td>162 [85%]</td>
</tr>
<tr>
<td>UGPA</td>
<td>3.6</td>
<td>3.6</td>
<td>3.6</td>
<td>3.5</td>
<td>3.7</td>
</tr>
</tbody>
</table>

The success of our program has exceeded our expectations. For the first 5 years of the program we had projected class sizes of 10-15 per year. We have exceeded these projections with class size during year one at 16 (13 fulltime and 3 part-time), class size during year 2 at 24 (22 fulltime), and the application rate for year 3 has doubled that of the previous year.

Our curriculum is patterned after similar masters programs (see Appendix B for specifics of our curriculum). During year one, students take core courses from three two-semester sequences: (a) probability and mathematical statistics; (b) applied data analysis; and (c)
biomedical disease process and communication. These sequences directly map to the 3 core competencies of analytical, biomedical, and communication and have been synchronized with regard to the schedule of topics. Students also take a programming class that teaches skills in the languages of R (the academic standard) and SAS (the industry standard). A qualifying examination is given in August before year 2.

Following their first year of coursework, many students are placed in paid or unpaid internships, with some working full-time during the summer and then part-time during the school year. During year two, students take a novel core course, Statistical Methods for Learning and Discovery, which surveys a number of techniques for high dimensional data analysis useful for data mining, machine learning and genomic applications. They also take a number of elective courses and work on a masters’ project (typically, either a thesis-style report of a methodological project directed by a B&B faculty member or a collaborative manuscript from an interdisciplinary project developed as part of their internship). They also take a practicum course focusing on job skills (e.g., developing their curriculum vitae and practicing job interview techniques).

An important component of our curriculum is its experiential learning opportunities, whereby students have the opportunity to practice (under supervision) serving as the statistician for an actual research project. The ideal such experiential learning opportunity is a paid internship that may also serve to provide the topic for a masters’ project. For our first class of students, paid internships outside Duke were provided by both FDA and Quintiles. Paid internships within Duke were provided by the Duke Cancer Institute, the Duke Biostatistics Core and the Duke Clinical Research Institute. Additionally, one student obtained part-time work with the National Evolutionary Synthesis Center at Duke. Feedback about student performance has been very positive.

Our first formal review (see Appendix C) of the masters program included an assessment of our curriculum and of student performance to date. The performance of our 13 full-time students has been excellent. Every student from our first class passed the qualifying examination. Two of these students out-competed PhD students from other programs for external internships at the FDA and at Quintiles and all full-time students were placed in internships, 10 of which are paid.

I.3.3. Existing synergistic programs

This proposal capitalizes on our overlap and synergy with two Duke PhD programs: the one in Computational Biology and Bioinformatics (CBB), which emphasizes skills in computation
and machine learning for modeling biological phenomena, primarily at the cellular level, and
that given in the Department of Statistical Science (DSS), home to the premier Bayesian
faculty in the world. DSS focuses on interdisciplinary methods and applications, but not
necessarily related to medicine. It provides a natural continuum and source for elective courses
pertaining to statistical methodology. Several of our faculty members serve on graduate
committees for these programs and we expect the reverse will be true. In addition, selected
courses from both the CBB and the DSS curriculum will be cross-listed in the Biostatistics
program and vice versa.

I.4. Integration of basic science and traditional models

Our program will incorporate an integrated learning approach, combining the analytic
substance of traditional didactic programs with the “laboratory” experience generally found in
the basic sciences and in the existing CBB program. This approach is consistent with the
nature of the discipline of biostatistics. The “traditional” component of didactic instruction
will allow students to obtain the theoretical and applied background to address multiple types
of problems, while the “laboratory” part of the curriculum will allow the student to develop an
appreciation for how the details of applying this background will differ from case to case.

II. Relationship to existing programs

II.1. Traditional programs in biostatistics

Most PhD programs in biostatistics reside in schools of public health or colleges of arts and
sciences. For those programs, the “natural laboratory” provided by a medical school
environment is missing or incompletely integrated. Linking our program with this “natural
laboratory” fosters close communication continuously during graduate training and will
provide the experiences required to fully achieve the three core competencies. In fact, our
graduate students will be encouraged to accompany physicians on rounds and in clinic, and
even in the surgical suite, as many of our faculty have done.

II.2. Goals of the proposed graduate program

Biostatisticians must master three core competencies: (a) analysis; (b) biomedical fluency; and
(c) communication. These competency goals were determined from interviews with successful
statisticians and their non-statistical collaborators, and are consistent with the literature.
The analytic competency is derived from the traditional emphasis of biostatistics, and focuses on both the theoretical and applied aspects of statistical science, statistical modeling, software development and application, and study design. This competency has been achieved when the biostatistician can confidently develop or select and apply appropriate analytical techniques to address the various research questions under study. This competency is primarily acquired through structured didactic instruction, work on the dissertation, and immersion into actual applications.

The biomedical competency requires a sufficient foundation in the biology of diseases and related sciences to be able to understand and appreciate both the principles and the nuances of statistical application to the problem at hand. In addition, current medical research relies on inter-disciplinary collaboration (Begg and Vaughn, 2012¹). While a biostatistician cannot be expected to become a scientific expert in each area under study, a fundamental level of background knowledge is essential. This competency is acquired not only through our novel didactic course in biomedical disease processes and communication, but also through immersion into the specific scientific milieus, where interdisciplinary interaction is routine.

The communication competency focuses on the unique role of the biostatistician within an increasingly complex and interdisciplinary research enterprise. This competency has been achieved when the biostatistician can effectively rephrase the scientific content into understandable statistical principles, and also embed the underlying biomedical context into the statistical models that are being applied or newly developed. This competency is approached in the first year through our novel course in which students are required to make presentations linking biomedical concepts with statistical thinking; communication is then reinforced and enhanced through affiliations actual faculty projects.

II.3. Uniqueness of the program in comparison with other biostatistics programs

Most programs in biostatistics focus heavily on the analytic competency. The biomedical and communication competencies are generally addressed through specific consulting projects or, in many cases, as on-the-job training. Because of the wealth of medical research opportunities at Duke and the close connections that have already been developed between B&B faculty members and various medical researchers, Duke provides a particularly fruitful environment for focusing on the biomedical competency.

¹ Begg M, Vaughan RD Are Biostatistics Students Prepared to Succeed in the Era of Interdisciplinary Science? (And How Will We Know?) The American Statistician Vol. 65, Iss. 2, 2011
A key feature of our program is exposure to the continuum of contemporary biomedical research. Our students will have basic knowledge in human biology and how the disruption of biologic systems leads to disease. This exposure will greatly facilitate their understanding of how key elements of biomedical systems are measured, leading to a better understanding of the data and better ability to respond to analytic issues presented by emerging technologies.

Our program will be noteworthy in that we focus on biomedical disease processes and communication competencies in addition to the standard analytic competency; and within the analytic competency, we are developing novel courses that will focus on state-of-the-science methodological issues likely to be encountered in 21st century biomedical research. Having an emphasis on all three competencies, supported by the world-class laboratory provided by the Duke biomedical enterprise, will make our program not only innovative, but also unique.

II.4. Comparison between our program and those of other programs at Duke and nationally

Faculty in the B&B Department focus specifically on problems that arise in biomedical research. In fact, most B&B faculty members have their offices located within the area that also houses the biomedical discipline that is their collaborative affiliation, such as cardiovascular disease or cancer. They develop specialized medical knowledge relative to this discipline and collaborate directly with Duke’s world class medical investigators. This “natural laboratory” in which our graduates will be trained is a distinguishing feature of our program.

In contrast, faculty in the Department of Statistical Science (DSS) at Duke, renowned for their expertise Bayesian statistics, develop methods and participate in applied research, but are not aligned with any specific partner discipline and their applications span the spectrum of statistical issues that arise in any research context, not necessarily medicine.

The other closely related program at Duke is the Computational Biology and Bioinformatics program, with which we share faculty and students. These scientists use computation for modeling fundamental biology, not necessarily related to medicine; biostatisticians, on the other hand, focus on quantifying and accounting for biological variability in making inferences. Our program will be complementary to both of these established programs.
II.5. Potential overlap with existing programs in the area

As noted above, our proposed program overlaps with and derives synergy from the two existing programs on campus. However, there are unique aspects of our program that distinguish it from the others. In particular, our program differs from Duke’s CBB program by its greater emphasis on systematic and didactic instruction and a focus on the inferential rather than the computational aspects of biological systems; the program differs from Duke’s Statistical Science program by its emphasis on methods and skills that derive from medical applications and its close affiliation with the school of medicine. Additionally, the program differs from UNC’s program in biostatistics because the Duke program is embedded in the Duke School of Medicine, rather than the School of Public Health, and our faculty are totally integrated into the “laboratory” of biomedical research; and the program differs from NCSU’s program in statistics in that it is focused on medical issues and NCSU does not have a medical school.

II.5.1. Duke Computational Biology and Bioinformatics (CBB)

The focus of the Biostatistics PhD program is both different from and complementary to the Duke graduate program in CBB offered through the Institute of Genome Sciences and Policy. While several B&B faculty teach and direct dissertation research in the CBB program, students in CBB orient more toward wet lab techniques and algorithmic approaches that are complementary to the inferential statistical approach. Fundamentally, although both programs deal with biomedical data, they are designed to explore different questions, using different methods.

The CBB program seeks to develop research scientists who are fluent in computational science, mathematics, and statistics, with more emphasis on training students through "on the job" training (3 full-semester laboratory rotations) than via methodology classes (only 4 courses are required). Whereas CBB students might attempt to understand the biological mechanisms underlying heart failure, they would be unlikely to design a clinical trial or to assess the comparative effectiveness of two different therapies for heart failure.

In contrast, the B&B PhD program emphasizes a strong methodological foundation and rigorous coursework (14 core courses are required in addition to electives, cognates, and RCR) in theoretical and applied biostatistics. Graduates from the B&B PhD program are expected to be skilled in developing new statistical methods and designs to analyze problems across multiple biomedical disciplines. While the CBB program is suited for students more oriented
toward computation and biology than to statistics, the B&B program is designed for students whose natural interest is in the analytical aspects of statistical theory and inference.

II.5.2. Statistical Science

The graduate program offered by the Department of Statistical Science (DSS) in the Trinity School of Arts and Sciences is a top 10 statistics program nationally, a national top 3 program in research productivity, and the premier Bayesian group in the world. As with B&B, DSS also has an emphasis on application and its students engage in collaborative, application-driven projects. Although some of their faculty members work with investigators in the medical school, medical applications are not their primary focus and they are not generally wholly embedded within the medical environment. Given the synergistic goals of the two programs, mutual benefit can be expected from sharing courses and teaching assistants.

II.5.3. Other Comparative Programs

We have specifically designed our core curriculum to be competitive with those of five of the top Biostatistics PhD programs in the country: UNC, Harvard, University of Washington, Johns Hopkins, and University of Michigan. The table in Appendix D outlines this comparison. As with our program, each of these programs emphasizes two years of theory, with slightly different nomenclature but similar course descriptions, and two years of basic methods/applications. Because many schools do not offer undergraduate majors in statistics or biostatistics, the first year PhD curriculum is generally coincident with the first year masters curriculum and more advanced students have the option to test out.

It should be made clear that this is one difference between such a program and the more “basic” programs that are essentially wet or dry lab based. Because biostatisticians spend a good deal of their time collaborating, they need a firm foundation in the topics covered in our core curriculum; they also require a sufficient theoretical background in order to select and engage in an independent research project. Unlike many other fields, our students do not enter the program with the “tools of the trade” and they are not likely to have the experience to select a topic area in their first year.

However, this situation should improve in time. As noted in a recent communication from the American Statistical Association titled “Interest in Statistics still rising:”

The number of high-school students who took the Advanced Placement (AP) Statistics Exam this year increased to nearly 170,000, a jump of 11% over the 2012 number. The percentage increase is in line with the average annual increases of 12% realized since
2003 and parallels the growth of AP Calculus AB. The strong growth in AP Statistics is considered one factor in the doubling of undergrad statistics degrees over the last five years.

As our program matures, we should be in a good position to select for students with an adequate background to enroll immediately in the second year classes.

Among PhD programs in Biostatistics, the UNC Biostatistics program is highly ranked and our closest competitor. Our proposed curriculum is very similar to theirs. However, we extend the curriculum in several ways: first, we have added our basic core sequence that focuses on biomedical concepts and communication; we are also developing a new course called “Survey in inferential theory and techniques” (described in section VI), designed to provide the graduate student with the latest techniques in order to keep up with the fast pace of the “big data” explosion. Most importantly, we are embedded within a medical center and our students will be working directly with medical investigators and will be introduced to the entire medical research enterprise. They will be aligned with faculty and research teams that meet regularly to deal with the issues that arise in the process of collaborative medical science. Because the UNC program is embedded within a school of public health rather than a school of medicine, its students are less likely to have the advantages of the “natural laboratory” described above.

The comparison in Appendix D demonstrates that our curriculum is very similar to that of UNC, with the addition of novel features mentioned above and described more fully below. Some programs allow more choice in the additional requisites, while our curriculum is more defined. We have also borrowed the Harvard model, reflected as well in some of the other programs, of requiring a cognate field such as epidemiology. The one substantive difference between our program and the others is that each of the other programs requires a course in epidemiology, while ours embeds epidemiological principles of study design and analysis into our core courses. A six-credit curriculum in epidemiology will be one of the possible cognates. Our program also requires two novel courses that will prepare our students for the big data explosion. In addition, while some programs require one or two semesters of statistical consulting (the Michigan program has a course titled Design and Analysis of Biostatistical Investigation that emphasizes the communication aspect), we require our unique biomedical concepts/communication sequence that is structured to familiarize students with a range of biomedical concepts to the extent that they are fluent communicators. This course prepares them for their second year, in which they engage in three two- to four-month research experiences. Additionally our students will participate in a weekly seminar course, with increasing responsibility as they progress through the program.
Other programs offered locally reside in the UNC and North Carolina State Statistics Departments. The former is a theory-based pure statistics department; the latter is a general statistics department that includes both theory and application orientations, but has no affiliation with a medical entity.

### III. Resources

#### III.1. Business Plan

In collaboration with the Graduate School, a budget has been developed and is attached as Appendix E. Briefly, we anticipate recruiting three to five PhD students each year along with approximately 20-21 masters students for a total incoming class of 25 and a steady state of about 20 PhD students after five years. We anticipate that first-year PhD students will be supported through institutional funds and that the Biostatistics Graduate Program will be responsible for second-year PhD students’ tuition remission and stipends. In the third through fifth year these students will be funded through individual faculty grants, traineeships, or graduate assistantships as described in Section III.4. mechanisms that are similar to other biostatistics programs. Other expenses for the entire program include effort for teaching, effort for a DGS, administrative effort, and an academic retreat. Revenue from approximately 42 masters students each year will contribute $1.5M to the Biostatistics Graduate Program. As noted in the Executive Summary, the budget supports PhD students in their second year, and provides a contingency of $200,000 for backstopping students beyond their third year.

- Tuition revenue from 42 students: approximately $1.5 Million
- Expenses – approximately $1.5 Million
  - Faculty Teaching (including advising): $443,000
  - PhD tuition, fees and stipends: $432,000
  - Director of Graduate Studies: $58,000
  - Masters Scholarships: $150,000
  - Administration: $71,000
  - Other operating: $159,000 (including retreat and G&A)
  - Space: $103,000

#### Contingencies

The financial plan maintains a program reserve of $1.1 Million. We expect no more than 12 students in any year to be in the 3rd to 5th years of study and thus in potential need of departmental backstop. As described below, because of our experience in placing masters interns and our success in attracting both methods and statistical coordinating center grants, we are confident that we will not need to rely on this backstop. In addition, to the extent that we do not access the backstop, it will accumulate.
Although we suspect that we will not be accessing this backstop, the Department has additional contingencies, as confirmed in the letter from the Dean of the School of Medicine, Nancy Andrews (Appendix I). Also, in addition to the many options for funding graduate students, one of our faculty members has obtained provisional assurances from several entities, including pharma and some foreign groups anxious to sponsor graduate students on fellowships. These commitments will not involve any obligation on either the student or the Department’s part.

**Course development**

Four higher-level courses will be required as core courses, three of which will be offered starting in the fall of 2015, as PhD students enter their second year. We have assurances from the former Chair of DSS, Alan Gelfand, that these three core courses are being taught by DSS and will be available to our students (see letter, Appendix A). The current Chair, Merlise Clyde, is in agreement with these plans. The fourth course, a novel survey course of modern inferential techniques and theory will be offered in the fall of 2016 and required of all PhD students and will be developed and taught by Dr. Zhiguo Li.

**Director of Graduate Studies**

A 25 percent effort for the DGS represents a realistic estimate of the time required to deal with the organizational and academic components of the program.

**Students**

Although at steady state we anticipate that the program will be admitting mostly advanced students who will test out of some of the first year core courses, our initial budgeting assumes that all students will begin with the year one core courses. Thus, we have assumed that students will take five years to complete the degree rather than four, and are conservatively assuming that all students will need five years of support.

We anticipate that after the first two years in the program each PhD student will be associated with an individual faculty grant or funded project, a traineeship, or a graduate assistantship that will cover the student stipend, as described below.
Inter-program communication

In discussions with the new Chair of the Department of Statistical Science (DSS), we agree that most top programs in statistics and biostatistics are in schools where those departments work closely together. DSS has created a world-renowned PhD program and they have indicated strong support for our proposed PhD program. Dr. Clyde and Dr. DeLong meet regularly to discuss curriculum sharing and best practices and will continue to do so. They will also connect the Directors of Graduate Studies for the two programs program and will encourage them to meet regularly as well. Moving forward, this guidance will be invaluable.

We also have a very strong relationship with the CBB program, as a number of our faculty members teach and supervise dissertations in this program. These faculty members have been and will continue to be a constructive force in shaping the current proposal.

Program evaluation

As suggested by the recent APC review, we anticipate a “review of the PhD program after three years (a formative review) and five years (a summative review) to assess: (a) faculty status and involvement; (b) grants and contracts where B&B faculty are lead PIs (distinct from being lead PIs of subcontracts); (c) student progress toward doctoral degrees (assessed early on by the measurements of the preliminary exam committees according to their evaluation rubrics); (d) self-directed research versus other-directed work by doctoral students; (e) at five years, publications and submissions in the field of biostatistics (or statistics) first-authored by doctoral students.

III.2. Resources available

III.2.1. Operational support

Financial management of the program will continue to be the responsibility of the Business Manager of the Department, assisted by the departmental Grants Administrators, Human Resources Manager, and Staff Assistant.

The current infrastructure that has been established for the Master of Biostatistics and the Clinical Research Training Programs will subsume the operational/management responsibilities for the Ph.D. program. Although the PhD program will incur additional administrative overhead, we benefit from the inherent economies of scale. The administrative and supervisory structure is as follows:
- The Master of Biostatistics Program Coordinator will assume the role of Director of Graduate Studies Assistant (DGSA), for the PhD program, managing the logistics required to matriculate students as appropriate within the standard operating procedures of both the School of Medicine and The Graduate School.
- The Assistant Director of CRTP will extend existing responsibilities for program databases, reporting systems, alumni follow-up, course and program evaluations, and learning management systems (e.g. Sakai) to include the PhD students as appropriate.
- The CRTP Program Coordinator will continue to work with the DGSA in supporting functions common to all three programs such as classroom scheduling, admissions and examining committee meetings, developing new operational procedures (and associated documentation) needed for the additional program, and assist in data entry for the internal databases.

III.2.2. Space

The B&B Department occupies 12,195 square feet on the 11th floor of Hock Plaza, where 20 faculty members have their offices. An additional 2,645 square feet on the ground level house 9 offices, 6 cubes and a conference room. In addition to the multi-functional CRTP classroom on the second floor, there are 2 classrooms with a capacity of 24 students each, and 3 conference/seminar rooms which can accommodate 10 people. Two team meeting rooms are equipped with large monitors so research or student groups can work together on projects, analyses, and presentations.

Two open seating areas provide informal meeting areas for students, faculty, and research staff. In addition, there are three break rooms and a copy/printer room. Workspace available for graduate students includes 13 desks in shared offices, 3 private study rooms, and a computer room which will accommodate 4 desktops for student use.

III.2.3. Computing

Computational and file servers are available for B&B students, faculty, and staff. These servers are housed in the Fitz East Data Center and are maintained by Net Friends in compliance with School of Medicine policies and procedures. The department provides computers and IT support for faculty and staff, including student and postdoc research associates. A computer workstation room will accommodate 3 to 4 desktops and a printer for student use.
III.3. Additional resources needed

Given the infrastructure in place with the two graduate programs being offered through the Biostatistics and Bioinformatics Department, the necessary resources are all available.

III.4. Potential or actual external funding

As described above and included in the budget, funding for students during the first two years of the program will be provided through central institutional and B&B departmental funds.

For years three through five, Biostatistics Departments traditionally offer several opportunities for funding graduate students, in part because the demand for biostatisticians who are trained at the masters level or higher is strong and likely to increase. Because of this increasing demand, fresh PhDs in Biostatistics generally move directly into government, industry, or faculty positions. There are very few who are willing to take a postdoc unless it is with a super star (one of our recently hired faculty members did a postdoc with Susan Murphy at Michigan; she just became a 2013MacArthur Fellow). Given the difficulty of advertising and assessing potential postdocs, selecting an appropriate graduate student, with whom we already have a relationship, is an obvious choice.

Many of our faculty, as well as faculty at other institutions, were trained as recipients of Graduate Research Assistantships, whereby they were funded to use and improve their skills in both biostatistics and communication by serving as members of a statistical team for a large collaborative study. Hence, although R01-type statistical research grants are relatively difficult to obtain, funding for collaborative work remains high and much of this work involves challenging biostatistical problems. Additionally, many of these collaborative projects are ongoing for several years and tend to have multiple statistical issues. Many of our graduate students will have the opportunity for (or will actually prefer) such funding, which represents training that is potentially complementary to the dissertation topic. In many cases the dissertation is derived from a problem that arises in the collaborative study.

This apprenticeship model is common in Biostatistics PhD programs. Many PhD biostatisticians working in both academia and industry were funded and trained through this mechanism, which has the advantage of embedding students in the actual environment where they will spend their careers. Because of the increasing demand for such quantitative skills, students are expected to (and generally eager to) perform at increasingly higher levels. The
DGS will monitor the student’s placement and progress to ensure the appropriateness of the student’s current biostatistical activities.

Thus, in the third through fifth years of the program, several models for funding will be available:

a. Some students might be funded through the traditional Teaching Assistantship mechanism that is common to graduate programs. In particular, the Department of Statistical Science (DSS) might fund some of our students for this purpose. Although each of our third-year PhD students will be committed to serving one semester as a Teaching Assistant with no compensation, DSS may have additional needs for Teaching Assistants. In discussions with the Chair of DSS, we have agreed to work on an equitable arrangement for supporting these students, should the need arise. We have not counted on this source of funding. Additionally, as our program evolves, we may be able to provide assistantships for our own masters level courses, rather than relying on supplementing the discretionary accounts of our instructors.

b. A second traditional means of graduate student support is through an R01 (or similar) grant on which a B&B faculty member is PI. As a result of a faculty survey, 10 of our faculty members anticipate being able to support a student on such funding and 6 additional faculty members can guarantee funding for a graduate student through a project in which they are involved, but are not the PI. The student will be supported to accomplish some portion of the stated goals of the grant. Of note, our faculty have only recently (since the beginning of the masters program) begun adding graduate student funding to grant applications.

c. A third potential source is a training grant, for which we will apply after we are able to demonstrate a track record.

d. As described above, Biostatistics programs are somewhat unique in that after two years of study, their students have training that is equivalent to that of newly trained masters biostatisticians, who are in great demand. Some students will be offered a Graduate Research Assistantship (GRA) to spend about 20 hours per week working collaboratively as a biostatistician on some medical research project(s). The GRA will be supervised by a B&B faculty member, who will ensure that the student is in a learning environment and is able to gain experience and expertise as he/she progresses through the program. This type of arrangement is typical of other Biostatistics PhD programs and our program has the advantage that our faculty members are embedded in the School of Medicine with access to multiple learning opportunities. Such opportunities will be available through the Duke Clinical Research Institute, the Duke Cancer Institute, the Duke Biostatistics Core, and other groups that employ biostatisticians (for example, NESCENT hired one of our masters students last summer – they might prefer a longer, graduated relationship). The type of experience will
include the range of biostatistical applications such as coordinating center projects (on which several of our faculty are PIs), clinical trials design and analysis, and observational data analysis. In some cases, the dissertation problem will be associated with the GRA project, but in other cases it could be independent.

e. Similar to the GRA, some of our students will obtain internships with industry. As mentioned above, one of our masters students competed successfully for an internship with Quintiles this past summer. In this case, the student will be supervised by both an industry PhD biostatistician and a B&B faculty member and the experience will be monitored to ensure progressively more challenging work.

f. We also anticipate a continued relationship with the FDA, where one of our masters students had an internship this past summer, under the direction of a B&B faculty member.

With only 4 PhD students entering per year, we anticipate no difficulty placing each student in one of the above positions. When the program is in its fifth year and at steady state with 12 PhD students in their third through fifth years of the program, our ability to secure additional funding will have matured to the point that we have developed relationships with a number of funding sources.

Although not necessary for the program’s financial viability, we are quite optimistic about the prospects for attracting additional outside financial support which, in turn, can be used for program expansion and enhancement. As an example, the masters’ program was able to obtain funding of $25,000 this school year from the Astellas USA Foundation, which we used to support an outstanding student from a historically disadvantaged minority group.

**Training grants**

The training grant mechanism is a likely funding mechanism for PhD graduate students in the B&B Biostatistics programs. A number of predoctoral training grants, such as the T32, exist through the NIH. Although we will pursue and investigate the T32 and others, the most appropriate and targeted is the National Institute of General Medical Science training grant. The website for this mechanism is: http://www.nigms.nih.gov/Training/InstPredoc/PredocDesc-Biostatistics.htm.

Our program should be well-positioned to attract an institutional traineeship from this branch of the NIH. As stated on their website:
“The purpose of these programs is to provide support for predoctoral training in biostatistical theory and evolving methodologies related to basic biomedical research including, but not limited to, bioinformatics, genetics, molecular biology, cellular processes, and physiology, as well as epidemiological, clinical, and behavioral studies. The goal is to ensure that a workforce of biostatisticians with a deep understanding of statistical theory and new methodologies is available to assume leadership roles related to the Nation’s biomedical, clinical, and behavioral research needs. Implementation will depend on the integration of biostatistics and basic biomedical sciences to create effective interdisciplinary training grant programs. The aim is to provide students with strong quantitative talents to pursue a wide range of opportunities in biostatistics research.

… Applications should address the challenges of melding two disparate cultures, statistics and biology, at both the faculty and student levels.”

Another grant specifically targeted toward training statisticians is the National Institute for Drug Addiction funded Research Education Grants for Statistical Training in the Genetics of Addiction (R25), which can be found described at: http://grants.nih.gov/grants/guide/pa-files/PAR-08-081.html. Again, this funding mechanism perfectly dovetails with our intended program, as stated on the website:

“This FOA invites applications focused on research education for the development and testing of new statistical models to address genetics-based research problems in addiction.”

In addition, the Integrative Graduate Education and Research Training (IGERT) program through NSF promotes interdisciplinary science and is another potential source of funding. Their website is http://www.nsf.gov/crssprgm/igert/gradopps.jsp.

We anticipate applying for at least one training grant during the first five years of the program. We have not included revenue from a training grant in our initial budget estimate.

Industry

The likelihood of external funding is especially high from participants in the pharmaceutical industry, who heavily rely upon PhD-trained statisticians and have internships designated specifically for them. Members of our Professional Advisory Council have commented that their organizations not only find the premises which underlie our educational program to be appealing, but also that they would be particularly likely to provide support for student
internships. The following links to internships for graduate students in biostatistics provide examples of the availability of such funding on a competitive basis:

http://www.amgen.com/careers/campus.html
http://magazine.amstat.org/blog/2012/12/01/annual-internship-listing/
http://www.quintiles.com/careers/global/career-paths/internships/
http://us.gsk.com/html/career/career-summer.html
http://careers.jnj.com/Internship-co-op-programs
http://www.mayoclinic.org/intern-biostats-rst/
http://www.nhlbi.nih.gov/funding/training/redbook/sibsweb.htm

Additionally, some members of our faculty have significant academic interactions with industry, such as presenting workshops on statistical methods. In particular, Dr. Shein-Chung Chow is a well-known author of several books on the design and analysis of pharmaceutical studies and is editor in chief of the Journal of Biopharmaceutical Statistics. He has approached several pharmaceutical companies and has obtained tentative commitments for fellowship support. The goal is to create an industrial affiliate program; members will donate funding in the form of student fellowships in the amount of $30,000 for three years, with no obligation on the part of the Department or the students who receive the fellowships.

III.5. Five year student, faculty and resources projections

III.5.1. Faculty

Biostatistics and Bioinformatics faculty number almost 50, of whom 6 are currently teaching the 7 core courses in the Masters in Biostatistics program; one is teaching two core courses and is training a junior faculty member to take one of them. Eight faculty members are teaching the second year masters program electives, two of which are being team taught by a senior faculty member and a junior faculty member.

The graduate program has additional potential teaching faculty drawn from the collaborating faculty from other departments, for example the Center for Human Genetics. Appointments of new teaching faculty to the Program will be made by the Chair of B&B, or his/her designee, on the basis of recommendations of the Graduate Faculty.
Graduate Faculty

Faculty members with long-term commitments to training in the program will be recommended by the Chair of the Department for membership in the Graduate Faculty. The table in Appendix F demonstrates that 30 of the 47 Biostatistics and Bioinformatics faculty members will be affiliated with the PhD program in various substantive roles. The letter T indicates that they are currently teaching courses in the Masters of Biostatistics program. Most of them will continue to teach and others, especially new faculty who have recently joined the department, will likely also teach as the program evolves. For example Zhiguo Li, who joined our faculty after doing a post-doc with Susan Murphy at the University of Michigan, will take responsibility for our novel course in statistical methods for learning and discovery and has already started developing it.

It is of note that, because the B&B Department has only recently initiated a teaching program, and many of our faculty members began their careers at Duke, most cannot list a significant history of mentoring PhD students. However, of the 30 faculty members involved in the program, 12 have been or are currently the primary advisor or co-advisor for a PhD dissertation and an additional 5 have served on dissertation committees. Another two faculty members in the Center for Human Genetics and an adjunct faculty member working part-time at the Durham VA have directed PhD dissertations and are anticipating doing so for the Biostatistics PhD. Many B&B faculty members are also supervising post-docs. All of these Duke faculty members are committed to taking on graduate students while several others, especially recent hires, are anxious to do so, as shown in the table (Role in program = A for being the primary advisor and C for willing to serve on a PhD committee).

As participants in the graduate program, graduate faculty are expected and eager to teach core and elective courses, contribute to curriculum development, serve as dissertation committee members and advisors for program students, and serve on committees associated with the program, including examination committees. Appointment to the graduate program faculty is made with this explicit understanding and expectation.

Faculty Appointments

Among the 20 faculty members who have at minimum served on a dissertation committee, only 7 are either tenured or on the tenure track and the others are active researchers who have been successful in attracting Duke, UNC, and NC State graduate students. Our non-tenure track faculty members have records of scholarship that are commensurate with those of tenure track faculty and they publish in the same journals as faculty of other institutions. In particular, Cliburn Chan is one of our most productive faculty members; he is not on the tenure track, but has a graduate student in Duke’s CBB program and has supervised 2 postdocs.
Issues regarding tenure and timelines in the School of Medicine require explanation. Tenure track appointments in the Basic Sciences Departments are on the seven year time clock; those in the Clinical Sciences are given the longer timeline (eleven year clock) because of their “clinical commitments,” which have an impact on their research productivity. The Biostatistics and Bioinformatics Department is one of two School of Medicine departments that are classified as both Clinical and Basic. However, B&B faculty members are generally hired in a model more similar to the clinical faculty model than the basic; they are mostly on a non-tenure track and are expected to spend a significant amount of time participating in research initiated by other investigators. Even tenure track faculty are not typically given the protected time or resources needed to establish an independent research career on the shorter timeline. Due to the collaborative nature of these appointments, our faculty members go through the Clinical Sciences APT pathway.

That said, it would be almost impossible to discriminate between our current tenure track and our non-tenure track faculty on the basis of research productivity and independent research. For this reason, and the fact that we would not be able to obtain the requisite number of tenure slots to sufficiently cover all of our current and future faculty members whose accomplishments and promise would warrant a tenure slot, the Biostatistics and Bioinformatics Department has been actively engaged in an initiative to find alternative mechanisms for ensuring a stable faculty, other than the tenure track. In order to enhance our recruitment efforts, we have obtained approval to hire new faculty members and to convert our more productive existing faculty members to extended horizon contracts. This model provides security for those who take the risks expected of an independent researcher, but avoids the additional risk to the University of a lifetime commitment. We are considering this situation our “tenure-equivalent.”

We have recently recruited two full professors on the extended track. Both have excellent track records as independent investigators and will be in our pool of potential thesis advisors. Their initial contracts are for five years; after three years, the contracts will convert to a rolling five-year contract which is renewed daily until notice of non-continuance, after which he still has a guaranteed five years. We are currently creating guidelines for determining which current faculty members qualify for these extended contracts. An additional 13 faculty members will most likely be converted to extended horizon within the next few years. In addition to our existing faculty and anticipated new hires, we should have no trouble matching PhD students with advisors.

Those faculty members who are likely to occupy our “tenure-equivalent” slots are active researchers who publish in the field and also attract external funding. Appendix K lists some
of the grants on which B&B faculty members serve as Principal Investigators. Many grants are “owned” by different entities, such as the Duke Cancer Institute, rather than the Department, so this tabulation is relatively crude and incomplete, but it provides a general idea of the ability to attract external funding. Additionally, the activity has steadily increased with more faculty submitting and being awarded grants; a number of grants (that would take considerable effort to tabulate) have been submitted and are pending review.

Note however, that PI status on externally funded grants is not always the best metric of B&B faculty scholarship and/or potential to direct student dissertations. Many of our faculty have leadership roles in large collaborative organization but are not the PIs on the primary funding mechanisms for these groups. We present Dr. Andrew Allen as an example. Though Dr Allen has been the PI of NIH funded methodology grants in the past (K25, R01, etc.), he is currently funded exclusively through collaborative research in which he has COI but not PI status. Never-the-less, these grants fund his postdocs and PhD students and his effort is almost exclusively focused on developing novel statistical methodology addressing problems that emerge from these collaborations.

Faculty Stability
It is worth noting that B&B has never lost a faculty member due to lack of funding. In fact, we have trouble meeting the demand for faculty biostatisticians, partly because of the overwhelming need and partly because of our philosophy of hiring onto our faculty only those with a track record of publications in the field and the potential to contribute to the stature of the department. We have also had success in negotiating with some of our funding partners, such as the Duke Clinical Research Institute, to underwrite protected time for methods research for our faculty who collaborate with them. Additionally, because of the stimulating research environment at Duke, turnover among our faculty is minimal. The table in Appendix F demonstrates these long-term commitments, as well as the number of relatively recent hires. Many of these faculty members have overlapping interests. We anticipate very little attrition among our graduate faculty and, in such case, should have no trouble finding an appropriate replacement for a dissertation advisor.

Student/faculty interactions
Our program is very similar to other PhD programs in biostatistics and our recruitment of students is also similar. Although the School of Medicine Basic Science departments, and possibly those in Arts and Sciences, recruit on an individual level with the intent of matching the background and specific interests of students to advisors who have funding and would like
to take on additional students, incoming biostatistics students are unlikely to have an 
undergraduate degree in the field or much familiarity with it. Students need an introduction to 
the various aspects of biostatistics through coursework and internships in order to have an 
adequate sense of the options.

To ensure consistent student-faculty interaction, we will pair each incoming PhD student with 
an initial research-oriented advisor/mentor who will meet regularly with the student and will 
track the student’s abilities and interests in order to help guide the student regarding the 
selection of a dissertation advisor. This initial mentor will serve as the student’s gateway to the 
research environment by having the student read some of his/her manuscripts, by introducing 
the student to other research-oriented faculty, and by including the student in his/her research 
meetings.

Students have multiple opportunities to interact with the graduate faculty, in addition to those 
who are teaching courses. One of our faculty members organizes and facilitates several 
student/faculty lunches during the school year, during each of which three or four faculty 
members briefly describe their research interests – followed by open discussion with the 
students. These sessions have been quite popular. PhD students will also be required to attend 
at least one of the ongoing journal clubs (Statistical Genetics, Bioinformatics, Clinical Trials, 
Causal Inference/Predictive Modeling). Additionally, the seminar course, “Problems in 
Biostatistics” will be attended by faculty members who will at times be presenting.

Further, students will be required to work closely with program faculty members during their 
two-to-four month research experiences starting at the end of the first year. These experiences 
will pair students with faculty engaged in methodological research and will allow students to 
learn more about the graduate faculty and their areas of expertise. Students are required to 
complete three of these experiences by the end of their second year. Note that this differs 
fundamentally from the master’s internship experience as there is an explicit pairing with a 
faculty member that could chair a dissertation and the focus is on the development of statistical 
methodology.

**Faculty Development**

The Department has had an established program for scientific development of faculty through 
regular seminars, identification of mentors in specific areas, and publication of technical 
papers. Pedagogical development is underway as well. Faculty teachers in the Master of 
Biostatistics program have been able to take advantage of the resources available within the 
Department through the Clinical Research Training Program (CRTP), as well as those of the 
Center for Instructional Technology of Duke University Libraries. In addition, the Department
engaged the services of a curriculum and faculty development consultant, Dr. Edward M. Neal, to work with faculty through group workshops and individual consultations on the development and evaluation of their courses. Dr. Neal is the former director of faculty development at the UNC Center for Teaching and Learning, and a regular instructor in Duke Graduate School’s Teaching IDEAS workshops. Workshop topics for the Master’s program faculty included Designing a College Course, Introduction to Active Learning, and Evaluating Communication Skills. The individual consultations focus on individual course planning, syllabus development, and evaluating student performance.

III.5.2. Administration

Director of Graduate Studies
The Director of Graduate Studies (DGS) is the official departmental or program administrator of the rules and regulations of the Graduate School, the designated advocate of the needs of the graduate program and graduate students, both within the department and in the University, and the initial advisor of all matriculating graduate students. Directors of Graduate Studies are nominated by the department chair or the program director and are appointed by the Dean of the Graduate School for a specified term of service.

Gregory Samsa, PhD has been the DGS for the Masters in Biostatistics program and he will continue in this role for the PhD program. He will devote 25% effort to this position.

III.5.3. Student job placement

Job placement efforts for the first cohort of students focus on building professional networks through internship experiences and student membership in the American Statistical Association (paid for by the Department), as well as formal didactic activities during BIOSTAT 701 Practicum Seminar. The specific aim of the didactic portion of BIOSTAT 701 is to integrate students’ internship experiences with guidance and practice in the communication skills required for obtaining professional employment upon graduation from the program. The course consists of 5 two-hour job preparation seminars scheduled to occur throughout the year - three during the fall semester and two during the spring semester. The course is led by the Program administrative team and includes activities involving human resources staff, department faculty, members of the program’s Professional Advisory Council, and career services consultants. The first three sessions included the following topics: Lessons Learned: Debriefing the Internship Experience; Immersion in the Field: Starting to Build Your Professional Network; The Recruiting Process; Looking, Applying, and Interviewing – Stories from the Trenches; and Preparing a Resume. The spring 2013 seminars will focus on preparing
for the interview, including practice through mock interviews. This course will be equally relevant for PhD students, who will be encouraged to attend in their final year of the program.

IV. Students

Our Masters in Biostatistics program has provided the infrastructure and track record for the extension to a PhD program. As indicated above, enrollment has far exceeded projections. Approximately two thirds of our applicants, offers of admission, and matriculants are international, and primarily Asian; about two thirds are women. We offered a full scholarship to one African American student in our first year and were able to attract her; she has proven to be the top student in that cohort. The GRE scores and undergraduate GPAs of our first two years of matriculants compare favorably with the other masters programs at Duke.

In a survey of admitted students who declined our offer, the following schools were listed when asked if they chose to accept another university’s offer of admission over ours:

- **Listed twice:**
  - Columbia University
  - UNC-Chapel Hill
  - Yale University

- **Listed once:**
  - Brown University
  - Georgetown University
  - Stanford University
  - SUNY Upstate Medical University
  - UC-Irvine
  - UC-San Diego
  - University of Minnesota

Entrance requirements for applicants to the PhD program will include more advanced level mathematics or statistics courses and these applicants will be evaluated more rigorously than those applying to the masters program. As with other biostatistics PhD programs, applicants will be required to have taken and received excellent grades in multivariate calculus and linear algebra. Real analysis is recommended. Test scores and letters of recommendation attesting to analytical aptitude are also important. Aside from a higher level of quantitative aptitude, we expect the sources and other characteristics of PhD applicants to be similar to those of our masters applicants.
We will continue to use a study section format for our admissions process, which will be stratified according to whether the applicant is applying to the Masters or the PhD program. Our admissions committee consists of all current teaching faculty and each applicant is evaluated by three reviewers who take into consideration the rigor of the student’s curriculum, the GPA, GRE scores, TOEFL scores where applicable, the writing quality and sentiment of the personal statement, and the recommendations. Dr. Elizabeth DeLong, chair of the B&B Department is the chair of the committee and is heavily invested in the success of the program; she will continue in this role until the program is at steady state.

As with other graduate programs, we will invite a select group of applicants for a visit so that there can be a mutual assessment of fit before accepting them. During this visit, they will learn more about our faculty and also potential research areas. We will also continue the practice of inviting students to a recruiting visit (and for PhD students, we will cover the cost of the visit). At that visit, they will have the opportunity to learn about the department and to meet some of the faculty members.

To maximize flexibility in attracting talented students with diverse academic backgrounds, students will enter the PhD program through three potential routes: direct acceptance of outstanding students, provisional acceptance into PhD through masters program with partial scholarship and re-assessment after the first year of course work and comprehensive exam; acceptance from the second year of our Masters program (we actually have four such students currently in our program who would like to continue).

Entering Students will be assigned an initial advisor, who will guide the student through selection of faculty affiliations for research-oriented experiences and will mentor the student regarding other issues related to the program.

**Differentiating the curriculum and experience of Masters and the PhD programs**

We have configured our program to optimize the experience of both sets of students in the following ways:

- Entering PhD students will each be paired with a research-oriented faculty mentor (not necessarily the eventual thesis advisor) who will ensure that the student is exposed to research in biostatistics.
The PhD core curriculum will be more demanding through the addition of two advanced level theory courses, an advanced course in theory of linear models, our novel survey course of modern inferential techniques and theory, and the progressive “Problems in Biostatistics” seminar.

Four currently offered elective courses will be required for the PhD, consistent with the requirements of other PhD programs.

By the end of the first year, PhD students will be required to select one or more research experiences by apprenticing for a period of time with faculty who have potential dissertation projects. In distinction, Masters students will be placed in internships for up to a full year in order to obtain practical exposure to the environment in which they intend to find employment.

PhD students will be required to advance through the seminar series to the point of presenting their contributions from their research experiences.

PhD students will be required to either teach an undergraduate course in the Department of Statistical Science or to serve as a teaching assistant under the supervision of a B&B faculty member for one or more semesters.

PhD students who, for whatever reason, decide to discontinue pursuit of the PhD, but want to obtain the masters degree, will be required to complete a six-credit masters project during which time they will be required to pay tuition similar to other masters degree students. In essence they will need to fulfill credit requirements for the masters degree.

PhD and masters students will gain from the associations with each other; advanced PhD students will serve as role models for the masters students and will also have opportunities to mentor and possibly tutor.

Both types of students will have some classes in other buildings besides Hock Plaza, with an opportunity to interact with other graduate students on campus. In particular, our students will have access to the new Trent Semans Center for Health Education, where graduate students and medical students gather to study and take courses.

**Diversity**

Our overall strategy for encouraging diversity is to become known as a program that proactively recruits gifted students from under-represented minority groups and provides every encouragement for them to succeed. We are working with Dr. Sherilynn J. Black, Director of the Office of Biomedical Graduate Diversity, to locate potential recruits. Also Dr. Gregory Samsa, the DGS of our Masters in Biostatistics program attends the annual Biomedical Research Conference for Minority Students (ABRCMS). Our first such student was unanimously nominated by the faculty as the most outstanding student among our first graduating class, and received a $25,000 annual scholarship from the Astellas USA Foundation.
V. Degree Requirements

V.1.1. Biostatistics theory and techniques

Any successful biostatistical researcher needs to be well versed in both statistical theory and biostatistical techniques. At the masters level, this requires a course sequence with enough breadth and depth so that the masters level biostatistician can correctly implement an appropriate analysis and interpret results. At the PhD level, a similar breadth of material is required but at a deeper level, so that the PhD level biostatistician is not only able to implement and interpret existing approaches but to modify/extend these techniques or to develop completely new methods for addressing emerging problems in biology and medical research.

At the same time we acknowledge that medical research is changing and that there will be a continuing emphasis on extremely large and complex datasets. Thus, in training biostatisticians for the future, our curriculum emphasizes not only traditional biostatistical theory and techniques (Table 1) but also modern inferential approaches that are applicable to emerging high dimensional problems (Table 2).

Table 1: Core courses that are typically found in biostatistics PhD programs

<table>
<thead>
<tr>
<th>Standard biostatistical methods and theory courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introductory Statistical theory (BIOSTAT 604)</td>
</tr>
<tr>
<td>Introductory Probability theory (BIOSTAT 601)</td>
</tr>
<tr>
<td>Advanced statistical inference (STA 732)</td>
</tr>
<tr>
<td>Probability and measure theory (STA 711)</td>
</tr>
<tr>
<td>Introductory biostatistical methods I (BIOSTAT 602)</td>
</tr>
<tr>
<td>Introductory biostatistical methods II (BIOSTAT 605)</td>
</tr>
<tr>
<td>Introductory statistical programming I (BIOSTAT 621)</td>
</tr>
<tr>
<td>Introductory statistical programming II (BIOSTAT 622)</td>
</tr>
<tr>
<td>Theory of linear models (STA 721)</td>
</tr>
<tr>
<td>Generalized linear models (BIOSTAT 719)</td>
</tr>
<tr>
<td>Categorical data analysis (BIOSTAT 714)</td>
</tr>
<tr>
<td>Survival analysis (BIOSTAT 713)</td>
</tr>
<tr>
<td>Analysis of correlated and longitudinal data (BIOSTAT 718)</td>
</tr>
</tbody>
</table>

Table 2: Novel core courses emphasizing modern inferential techniques and theory

<table>
<thead>
<tr>
<th>Novel biostatistical methods and theory courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical methods for learning and discovery (BIOSTAT 707)</td>
</tr>
<tr>
<td>Survey in modern inferential techniques and theory – de novo</td>
</tr>
</tbody>
</table>
Courses with numbers are all existing courses that are currently taught in either the Statistics PhD program or the Biostatistics masters program. Only the survey course in modern inferential techniques and theory will need to be developed de novo (see Section VI for a detailed description of this course).

Note that the technical depth of the material is at a level that meets or exceeds other top PhD programs in biostatistics. The survey course on modern inferential techniques and theory is novel and covers advanced material that is not covered in most PhD curricula. The theory and techniques covered in this course (empirical process theory, counting processes, semiparametric theory, etc.) are quite powerful and will give our graduates a significant advantage in solving difficult inferential problems during their dissertation research and beyond. For a summary of each of these courses, please see Section VI.0. For an overview of how these courses fit into the educational timeline of typical students, please see Appendix G.

V.1.2. Biology, Disease, and Medical Research

The best methodology research in biostatistics is that which is responsive to pressing inferential needs in medical research. Hence, it is important that the biostatistics researcher have a good handle on the issues that face the various medical disciplines with which they will interact. Having a firm grasp on the human biology of disease processes, and how various biologic quantities are measured is key. Without this background, it is difficult for the biostatistics researcher to translate the biomedical problem into the quantitative framework required to make inferential progress. In the traditional approach to biostatistics education, this material is deferred to on-the-job training and for many, never satisfactorily takes place. To address this deficit, we have developed a unique sequence of courses that provides an introduction to biology at a level suitable for practicing biostatisticians (Table 3). This course emphasizes the translation of the biomedical problem into a statistical framework. In this way, this course serves as a bridge between the biomedical content and the statistical techniques and theories learned in other courses.

V.1.3. Communication

Effective communication skills are essential across all areas of biostatistical research. Biostatisticians need to be effective communicators while formulating the problem that is to be addressed, devising and articulating a solution, and disseminating results. At the PhD level, they also need to be effective educators both to other members of a research team and to the next generation of biostatistical researchers. Unlike other biostatistics programs we explicitly
integrate the development of communication skills into our curriculum. Our novel core course sequence, *Introduction to the practice of biostatistics* (Table 3), explicitly emphasizes the communication aspect of translating biomedical problems into a statistical framework.

In addition, skill building exercises are emphasized in all courses; reports, presentations, and group exercises are present across the curriculum. Second, we will also have an ongoing seminar sequence that introduces students to current problems in biostatistics (see Section VI). This course will be required of all students throughout their tenure at Duke and their participation will evolve with their development. Thus, first year students will largely be attendees, while intermediate PhD students will be required to make presentations on areas related to their emerging research interests and advanced students will present their own methodologic research. The first seminar of each year will be devoted to “How to give a seminar” and will be presented by faculty and advanced graduate students. For student presentations, the final 10 minutes of the session will be dedicated to a constructive critique.

### Table 3: Novel core courses emphasizing communication

<table>
<thead>
<tr>
<th>Novel courses focusing on communication of statistical ideas in medical research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction to the practice of biostatistics I (BIOSTAT 603)</td>
</tr>
<tr>
<td>Introduction to the practice of biostatistics II (BIOSTAT 606)</td>
</tr>
<tr>
<td>Problems in biostatistics seminar</td>
</tr>
</tbody>
</table>

Note that BIOSTAT 603 & 606 address both biology and communication competencies.

### V.2. Elective courses

**Electives:** The program will offer at least 2 electives each academic year on a rotating basis. During the 2012-2013 academic year, we are offering: Statistical genetics and genetic epidemiology (BIOSTAT 710) and Observational studies (BIOSTAT 709). With the approval of the DGS, students can satisfy program requirements by taking courses from allied disciplines including those offered by: Department of Statistical Science, the Computational Biology and Bioinformatics Program, the Umbrella Program in Genetics and Genomics program, the Clinical Research Training Program, and the Global Health program. Students may also be able to take advantage of course offerings at regional universities, i.e., UNC and NCSU. A total of 5 elective courses are required.

**Cognate field electives:** Students will be required to explore in some depth a selected cognate field, a substantive field outside of biostatistics or statistics. Examples of cognate fields include
biophysics; environmental health; epidemiology; genetics; health policy and management; human development; molecular biology among others. The cognate field should be complementary to the student’s biostatistical area of interest. Two of the 5 total electives must be used to satisfy the cognate requirement and should form a coherent set of courses related to the cognate field selected. These courses should be primarily substantive, rather than quantitative, in nature and this requirement will likely be fulfilled in the third year, once the student has chosen a dissertation area.

Each student’s cognate program will be customized to the background and interests of the student. Primary faculty members in these supporting programs have expressed enthusiasm for collaborating with us to place students into appropriate sequences of courses (please see letters of support in Appendix A).

To help demonstrate how the cognate requirement could be fulfilled, we highlight three examples:

Student A. Student A’s methodology research focuses on a topic in statistical genetics. Student A has an undergraduate degree in mathematics but no genetics training beyond that found in an introductory biology class. Thus Student A would have a cognate program that would address that gap through courses in molecular and population genetics such as UPGEN 687: Complex traits and evolutionary genetics; MGM 532: Human Genetics.

Student B. Student B’s methodology research focuses on a topic concerning the analysis of observational data. For this research program, additional training in formal epidemiologic design in addition to topics in global health would form an ideal cognate program. For example, GLHLTH 702: Global health research: Design and Practice and GLHLTH 705: Global Health Research: Introduction to Epidemiologic Methods would give the student a strong background in epidemiologic design, measurement, and analysis.

Student C. Student C’s dissertation research focuses on analytic issues in clinical trials. Additional coursework in regulatory issues and research ethics would strongly complement this work and the career path the student is embarking on. For example, students might elect the following set of courses: CRP 253 Responsible Conduct of Research; CRP 254 Research Management; and CRP 262 Systematic Reviews and Meta Analysis.

We note that our program does not explicitly require students to take epidemiology courses as is the case in many (but not all) traditional biostatistics programs. Our approach was instead to imbed universally important epidemiologic principles, quantities, and design in the core
sequence courses and then give students the option to pursue further study in epidemiology through their cognate field requirement (this is entirely similar to Harvard’s approach) if it is complementary to the student’s area of biostatistical interest. For example, a student conducting dissertation research in statistical methods related to observational analyses would likely find further study in epidemiologic methods useful, while the student conducting research in proteomic analyses may find a cognate course sequence involving molecular biology and protein chemistry more useful.

V.3. Examinations and reasonable progress

**First year comprehensive exam:** This exam will be based on the first year of core coursework and is equivalent to the masters qualifying exam. It is typically offered in August, prior to the beginning of year 2, with one opportunity to remediate a failing grade. Incoming students who wish to test out of first year coursework will have to pass the theory and methods section of this exam. This exam is useful for counseling students as to the wisdom of pursuing the PhD degree, given that a default masters degree is not an option and will require a masters project. Hence, the second year of coursework could depend on the results of this exam.

**Second year qualifying exam:** This examination is based primarily on the second year of core coursework and represents a level of rigor beyond that required of masters students. All PhD students must pass the exam at a level sufficient to confirm the ability of the student in these core areas and the preparedness for progression to independent research. The theory exam is set, administered and examined by representatives of the Biostatistics Graduate Faculty. It is typically offered in May of year 2, with one opportunity to remediate a failing grade. Although we have structured our curriculum to be consistent with other similar programs, there is the possibility that a few PhD students will struggle with the typical second year course load. In that case, they will be able to delay some courses and consequently delay the qualifying exam, with permission.

**Preliminary Examination:** The PhD preliminary examination is administered in accordance with the guidelines set out by the Graduate School. Ordinarily students should pass the preliminary examination by the end of the third year. As part of this examination process, the candidate will submit a short written manuscript on the research area, to be presented and discussed at the exam. The manuscript will usually consist of a literature review and proposal for their anticipated area of thesis research. In addition to addressing the student’s readiness to
conduct independent research, the oral preliminary examination features questions and discussion designed to evaluate the competence and preparedness of the student in the broad field as represented by the Biostatistics courses during the preceding two years. Successful completion of the Exam qualifies the student as a PhD candidate in Biostatistics.

**Thesis Research and Defense:** Following successful completion of the Preliminary examination, students devote most of their time to original research under the supervision of their dissertation advisors and committees. They may also take courses for their Supporting Program, including advanced topics courses or courses outside Biostatistics and of immediate relevance to their thesis research topics. During the third year PhD students also begin to specialize, acquiring the tools and techniques required for their chosen area of expertise. The student, with counsel from the DGS, will select a primary dissertation advisor (if not selected earlier) and a dissertation committee. One member of the student’s committee will be either a clinical or a basic science investigator with expertise in subject matter relevant to the dissertation. A faculty member from the cognate field is also recommended. Since the program is tailored to the needs and career goals of each student, the exact program will differ for each student.

During the thesis research period, each student will have routine weekly interactions with the primary advisor, and will meet with the full thesis committee at least twice each semester to review progress, plans, thesis research scope and any other matters arising. Clear research-oriented metrics will be developed so that advisors (and committees) will be guided by uniform standards, both in training doctoral students and in evaluating their progress. The student will complete the thesis research, thesis writing and prepare for the thesis defense as prescribed by the regulations of the Duke Graduate School (available in full at the web site of the Graduate School). It is the expectation of the Biostatistics program that most students will complete their research and successfully defend the PhD within 4 or 5 years.

A student’s dissertation topic might or might not be derived directly from a project providing the student’s stipend. For example, a student might be funded through the Biostatistics Core that is managed by the B&B Department and the student might be affiliated with the lymphoma research team. However, the student may have chosen a dissertation advisor who is interested in problems related to censoring of data on toxicity. The student would develop methods for the research problem and might possibly use the lymphoma data to provide an example of the methodology.

Each student will be expected to publish or submit at least one first-author paper in the area of biostatistics or statistics, in addition to being included as a co-author on one or more papers.
with collaborators. PhD students will also be required to attend the Departmental bi-weekly mentoring session, where both senior and junior faculty members evaluate grant applications that are being prepared by B&B faculty members. For the student’s preliminary oral, we will adopt the convention used by the CBB program of requiring that it be in the form of a grant proposal, including an introduction, review of previous work, proposed work and justification regarding its relevance and likelihood of success,

**Satisfactory Progress:** Each year the student must make Satisfactory Progress toward the PhD degree. The Biostatistics Graduate Faculty meets annually to discuss the progress of each student. When there is a problem, the Director of Graduate Studies will write the student a letter summarizing the faculty's consensus and specifying any deficiencies (academic or otherwise) the student must correct, or any additional requirements the student must meet.

**V.4. Other requirements**

**Responsible Conduct of Research:** All basic medical science doctoral students are required to complete 18 hours of training in the Responsible Conduct of Research (RCR). Each biostatistics PhD student will be required to complete one of the following RCR training options to meet this Graduate School requirement:

- Attend GS 710A (12 hours, RCR Orientation at the Beaufort Marine Lab) prior to beginning graduate study and mandatory follow-up RCR training (4 hours, GS713) between Years 3 and 4 of their program; or
- Successful completion of CRP 253 Responsible Conduct of Research offered by the Clinical Research Training Program (22.5 didactic hours). (Please note that GS712 credit approval will be requested for CRP 253 per The Graduate School’s instructions as described in “Procedures to Request GS712 Credit Approval for Departmental RCR Training Events”)

**Seminars:** Throughout their participation in the Biostatistics Graduate Program, students are expected to attend B&B research seminars. In addition, they are required to actively participate in at least one of the Department journal clubs: Clinical Trials, Statistical Genetics, Bioinformatics, or Causal Inference. Students are also strongly encouraged to attend research seminars in the CBB, IGSP, Statistics and Decision Sciences and other allied programs as frequently as possible.

**Research-Oriented Experiences:** Our current masters program provides a mechanism whereby first-year students are introduced to three or four faculty members who describe their research and potential thesis projects during a lunchtime discussion. This introduction will
also serve to familiarize PhD students with the various faculty interests. After the first year of coursework, students will select one or more faculty members with whom to become affiliated to learn about research opportunities and also the practical aspects of a career in biostatistics. These experiences will serve to better acquaint the students with potential dissertation topics as well as funding opportunities.

**Teaching experience:** Each PhD student will be expected to be a teaching assistant for one semester in the Statistical Science Department.

A timeline for student progression through the program is presented in Appendix G.

**VI. New Course Offerings**

**Current problems in biostatistics (to be developed).** Advanced seminar on topics at the research frontiers in biostatistics. This course is comprised of readings of current biostatistical research and presentations by faculty and advanced students of current research in their area of specialization. Required by all students. Course instructor: rotating

**Advanced topics in modern inferential techniques and theory.** This course provides an introduction to topics that form the theoretical foundation of modern biostatistics and the inferential techniques for solving contemporary biostatistical problems. It covers topics in stochastic processes, weak convergence, semiparametric models and probability limit theory with an emphasis on those needed in biostatistics. Specific topics include stochastic processes, random walks, Markov chains, martingales, counting processes, Hilbert spaces for random vectors, semiparametric models, geometry of efficient score functions and efficient influence functions, the functional delta method, and the theory for M- and Z- estimators. Throughout, the presented theoretical results will be illustrated through application to real problems encountered during the faculty’s research. Additional topics including weak convergence and an overview of basic techniques in empirical process theory will be covered. This course provides students a ‘tool chest’ of advanced techniques for dealing with modern inferential problems. After completing this course, students will be well prepared for the inferential component of dissertation research. Prerequisites: Probability and measure theory (STAT 711) and statistical inference (STAT 732). Course instructor: Dr. Zhiguo Li
VII. Curriculum Vitae of Faculty Who Will Participate In the Program
Please see Appendix H.

VIII. Statement of Support from the Dean of the School of Medicine
Please see Appendix I.

IX. Student Learning Assessment Process Template
Please see Appendix J.
APPENDIX A

Letters of Support

- Department of Statistical Science
- PhD Program in Neurobiology
- PhD Program in Pharmacology and Cancer Biology
- PhD Program in Molecular Genetics and Microbiology
- Department of Molecular Genetics and Microbiology (Email from the Chair)
- University Program in Genetics and Genomics (2 letters)
- PhD Program in Immunology
- PhD Program in Cell Biology
- PhD Program in Molecular Cancer Biology
- PhD Program in Biochemistry
- PhD Program in Cell and Molecular Biology
- Masters Program in Global Health
- Clinical Research Training Program
- PhD Program in Computational Biology and Bioinformatics

(Letters start on next page)
The Graduate School  
Duke University  
Durham, NC

To whom it may concern,

This letter is a follow up to my letter of January 29, 2013 which is supplied below. I am comfortable with the response of the Department of Biostatistics and Bioinformatics to the recommendations of the Executive Committee of the Graduate Faculty and I remain fully supportive of this new program.

I am delighted to write on behalf of the Department of Statistical Science in support of the proposal by the Department of Biostatistics and Bioinformatics to establish a Ph.D. program in Biostatistics. We concur that the timing is right to go forward with this now. One of the conditions for a world-class biomedical research enterprise at Duke is a vibrant community of biostatisticians which, as a primary component, oversees a well-respected Ph.D. program. Furthermore, we have always felt that a strong research degree program in the Department of Biostatistics and Bioinformatics (B&B) would be an important complement to our outstanding Ph.D. program in Statistical Science. There is clearly a substantial demand for such Ph.D. programs and the rather unique flavor that is being proposed - with the attractive constellation of analytical skills, fundamental biomedical knowledge, and communication with regard to problem-framing and problem-solving - suggests that it will easily find a good share of the demand.

In terms of more explicit support from the Department of Statistical Science (DSS), three of our core courses (Advanced statistical inference, STA 732; Probability and measure theory, STA 711, and Theory of linear models, STA 721) are envisioned as part of the core curriculum for the Ph.D. in Biostatistics. We are committed to making space available in these courses to Ph.D. students from the B&B program who meet the prerequisites. Further, with regard to mutual benefit, we anticipate that some Ph.D. students in B&B will take other research courses that we offer and, reciprocally, that some students in our program might avail themselves of some of the special topics courses that will be offered through B&B.

In summary, we in DSS believe this new Ph.D. initiative has the potential to be a substantial boon
to B&B, to DSS, and to the University as a whole and we enthusiastically support its creation.

Sincerely,

[Signature]

Alan E. Gelfand  
Head of Department  
J B Duke Professor of Statistical Science and  
Professor, Environmental Sciences and Policy,  
Nicholas School
April 25, 2013

Elizabeth DeLong, PhD, MA
Professor and Chair
Department of Biostatistics and Bioinformatics
Duke Box 2721
Durham, NC 27710

Dear Dr. DeLong:

The Duke PhD program in Neurobiology is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be quite useful to the student.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As chair of the department and DGS for the Neurobiology program, we welcome your PhD students to take any of our courses for which they have the required background. We will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Best wishes.

Yours sincerely,

Stephen G. Lisberger, PhD
George Barth Geller Professor of Neurobiology
Chair, Department of Neurobiology
Investigator, Howard Hughes Medical Institute

Richard Mooney, PhD
George Barth Geller Professor of Neurobiology
Director of Graduate Studies
Dear Liz,

The Duke PhD program in Pharmacology and Cancer Biology is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be quite useful to the student.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As DGS of the Pharmacological Sciences program in the department of Pharmacology and Cancer Biology, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Jeffrey C. Rathmell, PhD.
Associate Professor
Department of Pharmacology and Cancer Biology
Director of Graduate Studies of Pharmacological Sciences
jeff.rathmell@duke.edu
April 28th, 2013

Elizabeth DeLong, Ph.D.
Professor and Chair,
Department of Biostatistics and Bioinformatics
Duke University School of Medicine

Dear Liz

The Duke PhD program in Molecular Genetics and Microbiology (MGM) is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We have been very impressed with your Master’s program students as they are a vibrant and interactive group of individuals. Taking this to the next level with a PhD program is the logical development of your successful training and education efforts.

I understand that you anticipate enrolling approximately 4 PhD students in any year and these will be distributed among several potential cognate programs. As Director for Graduate Studies for the MGM graduate program, I welcome your PhD students to take any of our courses for which they have the required background. These include courses in Gene Regulation, Critical Readings in Genetics, Human Genetics, Virology, Microbial Pathogenesis, and Advanced Topics in Infection Biology. We also believe that your requirement of a six credit outside of Biostatistics will be quite useful to both your students and those in our programs, as “wet-lab” PhD students will benefit from increased interactions with your trainees, who will bring a rigorous statistical perspective to genetic-centered research questions.

I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

I look forward to continuing to work together as your PhD training program develops and we continue to forge partnerships that build bridges across programs and departments to enrich the opportunities for everyone on campus.

Sincerely,

Raphael H. Valdivia
Associate Professor
Director for Graduate Studies,
Department of Molecular Genetics and Microbiology
Director, Duke Center for the Genomics of Microbial Systems
Dear Liz

From my perspective as chair of MGM, and former director of the Duke UPPG program from 2002-2009, including writing two competitive T32 renewals, I am delighted you are moving ahead with establishing a PhD program in your department at Duke. I was very impressed with your masters programs students whom I met at your holiday event, and they are a vibrant and interactive group of engaging and engaged students. Taking this to the next level with a PhD program is the logical development of your terrific training and education efforts and impact. Raphael will send a letter of support as DGS and director of the MGM graduate umbrella program, and my own view is also highly supportive that your students are welcome to register for and matriculate in any courses that they would find of interest for which we are responsible.

I look forward to continuing to work together as your program develops and we continue to forge partnerships that build bridges across programs and departments to enrich the opportunities for everyone on campus.

Best wishes

Joe

Joseph Heitman, MD, PhD

Chair, Department of Molecular Genetics and Microbiology James B. Duke Professor, Departments of Molecular Genetics and Microbiology (MGM), Pharmacology and Cancer Biology, and Medicine Duke University Medical Center
322 CARL Building Box 3546
Research Drive
Durham, NC 27710 USA

Tele: 919-684-2824
FAX: 919-684-5458
email: heitm001@duke.edu
Dr. Elizabeth DeLong  
Chair  
Department of Biostatistics and Bioinformatics  
Duke University School of Medicine  
Durham, NC 27710  

24 April 2013  

Dear Liz,  

The Duke University Program in Genetics and Genomics (UPGG) is eager to support your initiative to establish a much-needed PhD program in Biostatistics. We are also excited to learn that you will be requiring a six credit cognate program in an area outside of Biostatistics.  

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As Program Director of UPGG, I welcome your PhD students to take any of our courses for which they have the required background. Many of the courses in our program will be a natural fit for your students. I and the other leaders of our program (Beth Sullivan, Co-Director; Allison Ashley-Koch, Director of Graduate Studies) will be pleased to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses. We look forward to working together to enhance your new PhD program.  

Sincerely yours,  

Douglas A. Marchuk, Ph.D.  
Professor and Vice Chair; Department of Molecular Genetics and Microbiology  
Director: University Program in Genetics and Genomics.
April 25, 2013

Elizabeth DeLong  
Chair, Biostatistics and Bioinformatics Department  
Duke University  
2424 Erwin Rd.  
Durham, NC 27710

Dear Liz,

The Duke University Program in Genetics and Genomics is eager to support your initiative to establish a much-needed PhD program in Biostatistics. We are also excited to learn that you will be requiring a six credit cognate program in an area outside of Biostatistics. Genetics and genomics are a natural fit with biostatistics. As a genetic epidemiologist I am personally excited about this opportunity.

I understand that you anticipate enrolling no more than four PhD students in any year and these will be distributed among several potential cognate programs. As Director of Graduate Studies for the Program in Genetics and Genomics, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Allison Ashley-Koch, Ph.D  
Professor of Medicine  
Director of Graduate Studies  
University Program in Genetics and Genomics  
Duke University
April. 25, 2013

To: Professor Elizabeth DeLong
Chair
Department of Biostatistics and Bioinformatics
Duke University Medical Center, Box 2721

Dear Liz

The Duke PhD program in Immunology is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be beneficial to the biostatistics major.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As the Director of Graduate Studies for the Immunology program, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Yuan Zhuang, Ph.D.
April 29th, 2013

Elizabeth DeLong PhD
DCRI Associate Director
Co-Director, Outcomes Research and Assessment Group
Chair and Professor, Department of Biostatistics and Bioinformatics
Duke University Medical Center
Durham, NC 27710

Dear Liz:

As a follow-up to our discussions re: a PhD program in Biostatistics, I am writing to express our strong support and to say that the PhD program in Cell Biology fully and enthusiastically supports your efforts to establish a PhD program in Biostatistics. Given the nature of the discipline, we very much agree that a requirement for a six-hour credit cognate outside of Biostatistics will benefit the Biostatistics students. On that note, I serve as course organizer for CBI551/CMB551, Molecular Cell Biology - this course is our flagship cell and molecular biology course and is required of all Cell Biology graduate students as well as the graduate students admitted via the T32 Cell and Molecular Biology training grant. This would be an excellent course for your students; they will be exposed to a diversity of fundamental cell and molecular biology research areas as well as key literature in those fields – and will gain a strong appreciation for how, where and why biostatistics is playing an increasingly important role in basic research.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As DGS for the Cell Biology program, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Christopher Nicchitta, PhD
Director of Graduate Studies
Professor of Cell Biology and Biochemistry
Duke University Medical Center
April 25, 2013

Elizabeth Delong, PhD
Chair and Professor, Department of Biostatistics and
Bioinformatics, DUMC

Dear Liz,

The Duke PhD program in Molecular Cancer Biology is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be quite useful to the student.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As Director of Graduate Studies for the MCB program, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Best regards,

Ann Marie Pendergast, Ph.D.
Anthony R. Means Cancer Biology Professor
Vice Chair, Department of Pharmacology and Cancer Biology
Director of Graduate Studies, Molecular Cancer Biology
Duke University School of Medicine
P.O. Box 3813, Medical Center
Durham, NC 27710
April 25, 2013

Professor Elizabeth DeLong
Chair, Department of Biostatistics and Bioinformatics
Duke University

Dear Liz,

The Department of Biochemistry and the PhD program are very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We understand that the proposed new degree will require six credit hours outside of Biostatistics. I strongly believe that training in the biostatistics area needs to include this type of classroom and perhaps, in addition, even project experience. As professionals later, it is important that many if not most of PhDs from programs like yours be imbedded in experimental research groups, and this represents early substantive exposure to the scientific environment in which they will likely work.

I understand that you anticipate enrolling about 4 PhD students per year in this new program, and these students will be distributed among several potential cognate programs. As DGS for the Biochemistry program, I welcome your PhD students to take any of our courses for which they have the required background. I will be happy to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Leonard D. Spicer
Director of Graduate Studies
Department of Biochemistry
University Distinguished Service Professor
April 25, 2013

Elizabeth DeLong, PhD
Director, Biostatistics Core
Professor and Chair, Department of Biostatistics and Bioinformatics
Co-director, Outcomes Research and Assessment Group, Duke Clinical Research Institute
Duke University Medical Center
Durham, NC 27710

Dear Liz

The Duke PhD program in Cell and Molecular Biology is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be quite useful to the student.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As DGS for the CMB program, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Meta Kuehn, PhD
Director of Graduate Studies,
Cell and Molecular Biology Training Program
April 25, 2013

Dear Liz,

As you know, we at Global Health are very supportive of your proposal to offer a PhD program in Biostatistics. In fact, we have multiple collaborations with faculty in your department and understand the increasing demand for such quantitatively trained individuals.

We are also happy to know that you have proposed a cognate program in Global Health. Your students are certainly welcome to take any of our courses. I will be glad to work with you and Liz Turner, our primary collaborator in B&B, and students interested in Global Health to help select the most appropriate set of courses.

Sincerely,

Michael H. Merson, MD
Director, Duke Global Health Institute
Wolfgang Joklik Professor of Global Health
April 25, 2013

Elizabeth DeLong, PhD
Professor and Chair, Department of Biostatistics and Bioinformatics
Duke University Medical Center
Box 2721
Durham, NC 27710

Dear Dr. DeLong,

I am supportive of and enthusiastic about the proposed development of a PhD in Biostatistics program and the potential for collaborative endeavors between that program and the Clinical Research Training Program. Given that both programs will share the same departmental home, I believe there will some exciting cross-training opportunities for the programs to explore together.

As Director of the Clinical Research Training Program and a statistician myself, I am delighted that the proposed program will require coursework in a ‘cognate’ field complementary to, but outside of Biostatistics and believe this represents one of several particular strengths of the proposed program. There are multiple elective courses offered by our program (e.g., Research Management, Responsible Conduct of Research, Concepts in Comparative Effectiveness, Meta-analysis) that, I believe, will nicely complement the PhD curriculum and can be combined to serve as a ‘clinical research’ cognate program for interested PhD students. Students from the PhD program will be welcome to take any of our courses and I will be happy to work with the PhD program leadership and individual students to identify a series of courses that will provide the best fit for their interests.

Best Wishes,

Steven C. Grambow, PhD
Director, Clinical Research Training Program
Assistant Professor of Biostatistics and Bioinformatics
DUMC Box 2721
Durham, NC 27710
Email: steven.grambow@duke.edu
Work: (919) 684-1292
Dear Liz,

The Duke PhD program in Computational Biology and Bioinformatics is very happy to support your initiative to establish a much-needed PhD program in Biostatistics. We also believe that your requirement of a six credit cognate outside of Biostatistics will be quite useful to the student.

I understand that you anticipate enrolling no more than 4 PhD students in any year and these will be distributed among several potential cognate programs. As DGS for the CBB program, I welcome your PhD students to take any of our courses for which they have the required background. I will be glad to work with you and your individual students to ensure that they enroll in the most beneficial combination of courses.

Sincerely,

Scott C. Schmidler
Associate Professor of Statistical Science and Computer Science
Director of Graduate Studies, Computational Biology and Bioinformatics Program
Duke University
APPENDIX B

Curriculum Overview and Course Information

CURRICULUM OVERVIEW

Core Courses
Foundational courses required of all full-time Master of Biostatistics students.

BIOSTAT 601: Introduction to Statistical Theory and Methods I (3 Credits)
BIOSTAT 602: Applied Biostatistics Methods I (3 Credits)
BIOSTAT 603: Introduction to the Practice of Biostatistics I (3 Credits)
BIOSTAT 604: Introduction to Statistical Theory and Methods II (3 Credits)
BIOSTAT 605: Applied Biostatistical Methods II (3 Credits)
BIOSTAT 606: Introduction to the Practice of Biostatistics II (3 Credits)
BIOSTAT 707: Statistical Methods for Learning and Discovery (3 Credits)
BIOSTAT 621: Introduction to Statistical Programming I (1 Credit)
BIOSTAT 622: Introduction to Statistical Programming II (1 Credit)
BIOSTAT 701: Biostatistics Practicum (1-3 Credits)

Master's Project
Completed during the second year of study, all full-time Master of Biostatistics students will complete a master's project to demonstrate their mastery of biostatistics.

BIOSTAT 720: Master's Project (6 Credits)

Elective Courses
Full-time Master of Biostatistics students will select five elective courses during the second year of study.

BIOSTAT 708: Clinical Trial Design and Analysis (2 Credits)
BIOSTAT 709: Observational Studies (2 Credits)
BIOSTAT 710: Statistical Genetics and Genetic Epidemiology (2 Credits)
BIOSTAT 713: Survival Analysis (2 Credits)
BIOSTAT 714: Categorical Data Analysis (2 Credits)
BIOSTAT 718: Analysis of Correlated and Longitudinal Data (2 Credits)
BIOSTAT 719: Generalized Linear Models (2 Credits)

COURSE PLANNING

Fall Semester:
BIOSTAT 601: Introduction to Statistical Theory and Methods I (3 Credits)
BIOSTAT 602: Applied Biostatistical Methods I (3 Credits)
BIOSTAT 603: Introduction to the Practice of Biostatistics I (3 Credits)
BIOSTAT 621: Introduction to Statistical Programming I (1 Credit)
Spring Semester:
BIOSTAT 604: Introduction to Statistical Theory and Methods II (3 Credits)
BIOSTAT 605: Applied Biostatistical Methods II (3 Credits)
BIOSTAT 606: Introduction to the Practice of Biostatistics II (3 Credits)
BIOSTAT 622: Introduction to Statistical Programming II (1 Credit)

During the second year of study, full-time Master of Biostatistics students will typically take two core courses, a set of elective courses, and receive credit toward the completion of the master's project. A typical sequence is as follows:

Fall Semester:
BIOSTAT 707: Statistical Methods for Learning and Discovery (3 Credits)
BIOSTAT 720: Master's Project (3 Credits)
BIOSTAT 701: Biostatistics Practicum (1-3 Credits)
Elective Courses (4 Credits)

Spring Semester:
BIOSTAT 720: Master's Project (3 Credits)
BIOSTAT 701: Biostatistics Practicum (1-3 Credits)
Elective Courses (6 Credits)

AY 2012-2013 CORE COURSE DESCRIPTIONS

BIOSTAT 601: Introduction to Statistical Theory and Methods I
This course provides a formal introduction to the basic theory and methods of probability and statistics. It covers topics in probability theory with an emphasis on those needed in statistics, including probability and sample spaces, independence, conditional probability, random variables, parametric families of distributions, sampling distributions, and the central limit theorem. Core concepts are mastered through mathematical exploration, simulations, and linkage with the applied concepts studied in BIOSTAT 604.

Prerequisite(s): 2 semesters of calculus or its equivalent (multivariate calculus preferred). Familiarity with matrix algebra is helpful.
Corequisite(s): BIOSTAT 602, BIOSTAT 603
Credits: 3

Primary Instructor Name: Barry Moser
Semester Offered: Fall
Day(s) of Week Offered: Monday and Wednesday
Start/End Time: 10:00 AM – 11:15 AM

BIOSTAT 602: Applied Biostatistical Methods I
This course provides an introduction to study design, descriptive statistics, and analysis of statistical models with one or two predictor variables. Topics include principles of study design, basic study designs, descriptive statistics, sampling, contingency tables, one- and two-way analysis of variance, simple linear regression, and analysis of covariance. Both parametric and non-parametric techniques are explored. Core concepts are mastered through team-based case
studies and analysis of authentic research problems encountered by program faculty and demonstrated in practicum experiences in concert with BIOSTAT 603. Computational exercises will use the R and SAS packages.

Prerequisite(s): 2 semesters of calculus or its equivalent (multivariate calculus preferred). Familiarity with matrix algebra is helpful.
Corequisites(s): BIOSTAT 601, BIOSTAT 603, BIOSTAT 621
Credits: 3
Primary Instructor Name: Andrew Allen
Semester Offered: Fall
Day(s) of Week Offered: Monday and Wednesday
Start/End Time: 11:30 AM – 12:45 PM

**BIOSTAT 603: Introduction to the Practice of Biostatistics I**
This course provides an introduction to biology at a level suitable for practicing biostatisticians and directed practice in techniques of statistical collaboration and communication. With an emphasis on the connection between biomedical content and statistical approach, this course helps unify the statistical concepts and applications learned in BIOSTAT 601 and BIOSTAT 602. In addition to didactic sessions on biomedical issues, students are introduced to different areas of biostatistical practice at Duke University Medical Center. Biomedical topics are organized around the fundamental mechanisms of disease from both evolutionary and mechanistic perspectives, illustrated using examples from infectious disease, cancer and chronic/degenerative disease. In addition, students learn how to read and interpret research and clinical trial papers. Core concepts and skills re mastered through individual reading and class discussion of selected biomedical papers, team-based case studies and practical sessions introducing the art of collaborative statistics.

Corequisite(s): BIOSTAT 601, BIOSTAT 602
Credits: 3
Primary Instructor Name: Cliburn Chan
Secondary Instructor Name: Gregory Samsa
Semester Offered: Fall
Day(s) of Week Offered: Tuesday and Thursday
Start/End Time: 11:45 AM – 1:00 PM

**BIOSTAT 604: Introduction to Statistical Theory and Methods II**
This course provides formal introduction to the basic theory and methods of probability and statistics. It covers topics in statistical inference, including classical and Bayesian methods, and statistical models for discrete, continuous and categorical outcomes. Core concepts are mastered through mathematical exploration, simulations, and linkage with the applied concepts studied in BIOSTAT 605.

Prerequisite(s): BIOSTAT 601 or its equivalent
Corequisite(s): BIOSTAT 605, BIOSTAT 606
Credits: 3
Primary Instructor Name: Kouros Owzar
Semester Offered: Spring
Day(s) of Week Offered: Tuesday and Thursday
Start/End Time: 10:30 AM – 11:45 AM

**BIOSTAT 605: Applied Biostatistical Methods II**
This course provides an introduction to study design, descriptive statistics, an analysis of statistical models with continuous, dichotomous and survival outcomes, with one or more predictor variables. Topics include mixed effects models, likelihood and Bayesian estimation, generalized linear models (GLM) including binary, multinomial and log-linear models, basic models for survival analysis and regression models for censored survival data, clustered data, and model assessment, validation and prediction. Both parametric and non-parametric techniques are explored. Core concepts are mastered through team-based case study and analysis of authentic research problems encountered by program faculty and demonstrated in practicum experiences in concert with BIOSTAT 606. Computational exercises use the R and SAS packages.

Prerequisite(s): BIOSTAT 602 or its equivalent
Corequisite(s): BIOSTAT 604, BIOSTAT 606, BIOSTAT 622
Credits: 3

Primary Instructor Name: Hussein Al-Khalidi
Semester Offered: Spring
Day(s) of Week Offered: Monday and Wednesday
Start/End Time: 2:00 PM – 3:15 PM

**BIOSTAT 606: Introduction to the Practice of Biostatistics II**
This course revisits the topics covered in BIOSTAT 603 in the context of high-throughput, high-dimensional studies such as genomics and transcriptomics. The course will be based on reading of both the textbook and research papers. Students will learn the biology and technology underlying the generation of “big data”, and the computational and statistical challenges associated with the analysis of such data sets. As with BIOSTAT 603, there will be strong emphasis on the development of communication skills via written and oral presentations.

Prerequisite(s): BIOSTAT 603
Corequisite(s): BIOSTAT 604, BIOSTAT 605
Credits: 3

Primary Instructor Name: Cliburn Chan
Secondary Instructor Name: Gregory Samsa
Semester Offered: Spring
Day(s) of Week Offered: Tuesday and Thursday
Start/End Time: 8:30 AM – 9:45 AM
**BIOSTAT 707: Statistical Methods for Learning and Discovery**
This course surveys a number of techniques for high dimensional data analysis useful for data mining, machine learning and genomic applications, among others. Topics include principal and independent component analysis, multidimensional scaling, tree based classifiers, clustering techniques, support vector machines and networks, and techniques for model validation. Core concepts are mastered through the analysis and interpretation of several actual high dimensional genomics datasets.

Prerequisite(s): BIOSTAT 601 through BIOSTAT 606, or their equivalents
Credits: 3
Primary Instructor Name: Kouros Owzar
Semester Offered: Fall
Day(s) of Week Offered: Tuesday and Thursday
Start/End Time: 10:30 AM – 11:45 AM

**BIOSTAT 708: Clinical Trial Design and Analysis**
Topics include early phase through late phase clinical trials, including two-stage, Simon’s optimal design, parallel group, crossover, cluster randomized, and adaptive designs. Objectives such as endpoint selection, dose range, maximum tolerated dose, non-inferiority, surrogate outcomes, and safety will be considered. Methods for group sequential testing, will include fixed group sequential, O’Brien-Fleming, Pocock, one-sided, Tsiatis, Whitehead triangular and other tests. Wang method, repeated confidence intervals, and a range of related topics in monitoring trials.

Prerequisite(s): BIOSTAT 601 and BIOSTAT 604, or permission of the Director of Graduate Studies
Credits: 2
Primary Instructor Name: Bercedis Peterson
Semester Offered: Spring
Day(s) of Week Offered: TBA
Start/End Time: TBA

**BIOSTAT 709: Observational Studies**
Methods for casual inference, including confounding and selection bias in observational or quasi-experimental research designs, propensity score methodology, instrumental variables, and methods for non-compliance in randomized clinical trials.

Prerequisite(s): BIOSTAT 601 and BIOSTAT 602, or permission of the Director of Graduate Studies
Credits: 2
Primary Instructor Name: Kevin Anstrom
Secondary Instructor Name: Carl Pieper
BIOSTAT 710: Statistical Genetics and Genetic Epidemiology
Topics from current and classical methods for assessing familiality and heritability, linkage analysis of Mendelian and complex traits, family-based and population-based association studies, genetic heterogeneity, epistasis, and gene-environmental interactions. Computational methods and applications in current research areas. The course will include a simple overview of genetic data, terminology, and essential population genetic results. Topics will include sampling designs in human genetics, gene frequency estimation, segregation analysis, linkage analysis, tests of association, and detection of errors in genetic data.

Prerequisite(s): BIOSTAT 601 and BIOSTAT 604, or permission of the Director of Graduate Studies
Credits: 2

Primary Instructor Name: Andrew Allen
Semester Offered: Spring
Day(s) of Week Offered: TBA
Start/End Time: TBA

BIOSTAT 713: Survival Analysis
Introduction to concepts and techniques used in the analysis of time to event data, including censoring, hazard rates, estimation of survival curves, regression techniques, applications to clinical trials. Interval censoring, informative censoring, competing risks, multiple events and multiple endpoints, time dependent covariates; nonparametric and semi-parametric methods.

Prerequisite(s): BIOSTAT 601 and BIOSTAT 604, or permission of the Director of Graduate Studies
Credits: 2

Primary Instructor Name: Yuliya Lokhnygina
Semester Offered: Fall
Day(s) of Week Offered: Tuesday and Thursday
Start/End Time: 3:00 PM – 4:15 PM

BIOSTAT 714: Categorical Data Analysis
Topics in categorical modeling and data analysis/contingency tables; measures of association and testing; logistic regression; log-linear models; computational methods including iterative proportional fitting; models for sparse data; Poisson regression; models for ordinal categorical data, and longitudinal analysis.

Prerequisite(s): BIOSTAT 601, BIOSTAT 602, BIOSTAT 604, and BIOSTAT 605, or permission of the Director of Graduate Studies
BIOSTAT 718: Analysis of Correlated and Longitudinal Data
Topics include linear and nonlinear mixed models; generalized estimating equations; subject specific versus population average interpretation; and hierarchical model.

Prerequisite(s): BIOSTAT 601, BIOSTAT 602, BIOSTAT 604, and BIOSTAT 605, or permission of the Director of Graduate Studies
Credits: 2

BIOSTAT 719: Generalized Linear Models
The class introduces the concept of exponential family of distributions and link function, and their use in generalizing the standard linear regression to accommodate various outcome types. Theoretical framework will be presented but detailed practical analyses will be performed as well, including logistic regression and Poisson regression with extensions. Majority of the course will deal with the independent observations framework. However, there will be substantial discussion of longitudinal/clustered data where correlations within clusters are expected. To deal with such data the Generalized Estimating Equations and the Generalized Linear Mixed models will be introduced. An introduction to a Bayesian analysis approach will be presented, time permitting.

Prerequisite(s): BIOSTAT 601, BIOSTAT 602, BIOSTAT 604, and BIOSTAT 605, or permission of the Director of Graduate Studies
Credits: 2

BIOSTAT 720: Master’s Project
Completed during a student’s final year of study, the master’s project is performed under the direction of a faculty mentor and is intended to demonstrate general mastery of biostatistical practice.
Prerequisite(s): BIOSTAT 601 through BIOSTAT 606
Corequisite(s): BIOSTAT 607
Credits: 6

**BIOSTAT 621: Introduction to Statistical Programming I**
This class is an introduction to programming, targeted at statistics majors with minimal programming knowledge, which will give them the skills to grasp how statistical software works, tweak it to suit their needs, recombine existing pieces of code, and when needed create their own programs. Students will learn the core of ideas of programming — functions, objects, data structures, input and output, debugging, and logical design — through writing code to assist in numerical and graphical statistical analyses. Students will learn how to write maintainable code, and to test code for correctness. They will then learn how to set up stochastic simulations and how to work with and filter large data sets. Since code is also an important form of communication among scientists, students will learn how to comment and organize code to achieve reproducibility. The class will be taught in both the SAS and R language. Programming techniques and their application will be closely connected with the methods and examples presented in the co-requisite courses.

Corequisite(s): BIOSTAT 602
Credits: 1

Primary Instructor Name: Megan Neely
Semester Offered: Fall
Day(s) of Week Offered: Thursday
Start/End Time: 9:00 AM – 10:15 AM

**BIOSTAT 622: Introduction to Statistical Programming II**
This class is an introduction to programming, targeted at statistics majors with minimal programming knowledge, which will give them the skills to grasp how statistical software works, tweak it to suit their needs, recombine existing pieces of code, and when needed create their own programs. Students will learn the core of ideas of programming — functions, objects, data structures, input and output, debugging, and logical design — through writing code to assist in numerical and graphical statistical analyses. Students will learn how to write maintainable code, and to test code for correctness. They will then learn how to set up stochastic simulations and how to work with and filter large data sets. Since code is also an important form of communication among scientists, students will learn how to comment and organize code to achieve reproducibility. The class will be taught in both the SAS and R language. Programming techniques and their application will be closely connected with the methods and examples presented in the co-requisite courses.

Corequisite(s): BIOSTAT 605
Credits: 1

Primary Instructor Name: Megan Neely
Semester Offered: Spring
Day(s) of Week Offered: Monday
BIOSTAT 701: Biostatistics Practicum
To provide an opportunity to gain practical experience, using acquired competence in statistical analysis, the fundamentals of biology, and the skills of professional communication. This opportunity is used to assess and guide student initiative, leadership, decision-making, and accountability outside the classroom environment. These objectives can be met by an internship or another mechanism approved by the Director of Graduate Studies. This practicum may be continued through successive semesters, earning one credit per semester for up to three credits (credits may not be applied toward required elective credits). Credit amount is contingent upon approval by the Director of Graduate Studies as well as successful completion of the practicum experience.

Prerequisite(s): BIOSTAT 601 through BIOSTAT 606
Credits: 1-3
APPENDIX C

Formal review of the entirety of the
Masters in Biostatistics program

November 12, 2012

Note: The curriculum review is limited to the first-year courses, as these are the courses that are being offered for a second, rather than a first, time. For the purposes of this review, the courses are being considered as sequences (e.g., the programming sequence). Its content is based on discussions with Ed Neal PhD, our curriculum consultant.

The reasons that we are evaluating the curriculum include the following:

- Our educational program should be consistent with its advertised program goals
- Our educational program should have a strong likelihood of producing the desired student outcomes
- We need to coordinate the MS and proposed PhD programs – for example, we need to decide whether our first-year MS courses will also be appropriate for PhD students

Our metrics for success include the following:

- Alignment of the curriculum with desired program goals
- Alignment of the curriculum with desired teaching methods
- Alignment of the curriculum with desired student outcomes

How the evaluation has been performed:

- Interviews with instructors
- Review of student performance (e.g., grades, qualifying examination)
- Student feedback

Note: Our program philosophy is one of active-learning-based education. Often, active-learning-based education de-emphasizes rote learning of facts through traditional lecture, and emphasizes active engagement with the material through simulations, group projects, applied problem sets, reports and discussion. Instructors replace “What materials should I cover?” with “What tasks do I expect my students should be able to perform?”
Narrative (organized by course sequence):

Probability and mathematical statistics sequence (BIOS201&204):

PhD students in the statistics department are responsible for two sequences of courses on statistical inference. (For the present purposes, a statistical inference sequence is the same as a probability and mathematical statistics sequence.) The first sequence is basic but reasonably thorough, and is taken as an advanced undergraduate. The second sequence is more advanced (e.g., taught from a measure-theoretic perspective), and is usually taken in the first year of graduate school.

Our first-year core sequence is roughly equivalent to the advanced undergraduate sequence taken by students in the statistics department. It is sufficient for students having a terminal masters’ degree. It would also adequately prepare our second-year doctoral students to take the first-year inference sequence in the Department of Statistics. Our doctoral students would take that sequence of courses during their second year.

In general, it appears that our current first-year inference sequence won’t require significant modification to deliver to doctoral students. Possible modifications are a modest increase in the level of difficulty of the problem sets, and perhaps by moving up one level in textbook (e.g., from Bain & Englehart to Casella & Berger). Separate curricula, separate examinations, and other more substantial changes won’t be required.

On the surface, both BIOS201 and BIOS204 are traditional lecture-based courses, covering a standard sequence of material at the level of Bain & Englehart. For example, BIOS201 covers the first 6 chapters of Bain & Englehart, including topics such as probability, random variables, distributions, expectation, moment generating functions and the distribution of functions of random variables.

In reality, however, there is considerable congruence between the teaching methods used in this course sequence and the principles of active-learning-based education. Most notably, both courses are heavily problem-based. Moreover, the assignments and testing are organized in a way that strongly encourages students to keep up with working the problems, and to provide immediate feedback on performance. In
BIOS201, Dr. Moser gives weekly assignments (typically, 10 problems from the end of the chapter) and provides meticulous comments. (He also gives 2 mid-term examinations and a final examination). Dr. Moser also provides practice in how to frame problems – often, using examples from clinical trials.

In BIOS204, Dr. Owzar also assigns numerous problem sets. Lecture notes are provided, and are viewed as a way of working around some of the deficiencies of the text (Bain & Englehart). Although a solid mathematical foundation is assumed (e.g., students need to be familiar with epsilon-delta arguments and definitions of limits), BIOS204 is also designed things in a way to make remediation possible.

The course organization of BIOS204 is designed to ensure that students don’t fall behind. Quizzes are provided soon after material is encountered, and much of it is verbatim from what is covered in class. Indeed, including more problems that require applying the course content would probably be helpful, as this could provide a better picture of the facility that students actually have with the material. The grading scale equally weights homework, quizzes, technical reports, a midterm exam and a final exam (these exams being approximately weighted as 60% on previously encountered material, 20% on straightforward extensions of that material, and 20% on more challenging problems).

The instructional method is noteworthy in that the assignments include are two 5-page technical reports on a statistical topic (e.g., the EM algorithm), including an introduction, methods, results, and discussion. These provide directed practice in communicating fundamental concepts of statistical methodology, and would also provide experience in creating the technical appendices that often supplement manuscripts on more applied topics. Such technical reports are unusual for inference courses, a wonderful idea, and something from which other active-learning-based exercises might be based.

**Applied data analysis sequence** (BIOS202 & 205):

The applied data analysis sequence has a heavy emphasis on comprehension through active exposure to the course materials. The topics are relatively standard. In BIOS202, Dr. Allen starts from square 1 and covers the usual content up to multi-predictor models, for example: central tendency; measures of association; distinction between population and sample; 1- and 2-sample estimation and testing in parametric
and non-parametric contexts; the central limit theorem; analysis of count data using contingency tables and Poisson regression; simple linear regression; and Kaplan-Meier curves and the log-rank test. There is a heavy emphasis on developing an intuitive motivation for the statistical methods in question, and derivations are included. Almost all the methods are illustrated in both R and SAS. Simulation is used to help develop intuition – examples of simulation exercises include a bootstrapping assignment and a permutation test assignment.

Grades are assigned based on weights of 50% weekly quizzes, 30% homework, and 20% final exam. The homework assignments are intended to also offer practice in written communication skills – for example, in producing project reports with introduction, methods, results and conclusions sections. Lecture notes are used rather than a textbook. For the homework assignments students are encouraged to work in groups, but must write up the results on their own.

As with many of the other core courses, the regimented nature of the quizzes and homework assignments is intended to maintain student engagement, and also to identify problems with student performance as early as possible.

In BIOS205, Dr. Al-Khalidi follows roughly the same philosophy. Grades are assigned based on weights of 20% for weekly quizzes, 25% for the mid-term exam, 15% for homework, and 40% for the final exam. The instructor tends to be lenient in grading the quizzes – getting something wrong once isn’t a problem but getting something wrong twice is. Students are encouraged to work in groups but write things up on their own. The homework has a standard reporting format (e.g., including a description of the analyses, the results of the analyses, and the interpretation of the results). SAS is used mostly but not entirely. Lecture notes are used rather than a textbook.

Course content is organized around a full-rank matrix-based perspective on linear models, and includes multiple regression, variable selection, 1-way analysis of variance (from a cell viewpoint), 2-way analysis of variance, analysis of covariance, and GLIM (e.g., including logistic regression and proportional odds). There is considerable emphasis on using matrix algebra to develop intuition for simple designs, which is then extended forward to more complicated cases.
PhD students entering with a MS in statistics (or similar training) could probably place out of the applied analysis sequence – otherwise, given its emphasis on comprehension and communication through technical reports, it would probably be valuable for PhD students to take the courses with little or no modification. One idea (equally applicable to the inference sequence) would be to have the first-year masters’ qualifying exam also serve as a competency test for doctoral students that wish to place out of first-year courses. Doctoral training should include one additional linear models courses, perhaps taught at a greater level of methodological sophistication than our current second-year electives.

**Biology and communication sequence (BIOS203 & 206):**

A draft manuscript provides additional detail about this course sequence. Briefly, it is a 3-credit course, taught in two 75-minute sessions per week. The main goals are to teach students to:

- Discuss biology with investigators
- Discuss statistics with investigators
- Finally, to integrate the two

The biological content combines a moderate level of technical sophistication (e.g., detailed knowledge of biochemistry is not required when studying various biological pathways) with a deeper exposure to unifying concepts such as evolution. This training is intended to, among others: (a) put the student in a position to read the medical literature with confidence; and (b) effectively discuss biomedical content with biomedical researchers, statistical colleagues, and others. The statistical content is a systematic treatment of the “art” of statistics – for example, how to design a study, how to formulate an analysis plan, how to interview an investigator, etc.

The presentation of the course content is very much in the style of active-learning based education. Many of the assignments – for example, to formulate responses to the study questions that accompany the day’s readings – are first performed individually but processed at the level of the group. Much of the biological content is transmitted by the students in 10-minute PowerPoint presentations. Much of the statistical content involves writing followed by group discussion.

In response to previous student comments, the organization of the class is much more systematic than last year. Also, more of the assignments have been shifted from the
level of the entire class to the level of the work group. The instructors have been quite encouraged by the quality of the student presentations, and also from comments from the previous class of students about how the course content was helpful during their internships.

This course sequence can probably be maintained, as is, for doctoral students.

**Programming sequence (BIOS221 & 222):**

This is currently a 1-credit course sequence, taught for 1 hour per week. There are 3 main goals.

1. To teach basic proficiency in SAS and R, these skills being key to successful employment after graduation (and also a fruitful internship experience)
2. To teach basic programming ideas (e.g., modularization)
3. To support the applied data analysis sequence by providing just-in-time information about programming techniques for assignments (e.g., how to perform a bootstrap simulation exploring the robustness of a regression model)

The primary instructional technique is lecture/demonstration during class time, with self-study assignments of increasing difficulty used as homework. Group collaboration is encouraged. The instructional technique is somewhat based on principles of active learning, mostly because of the high level of engagement that students (and groups of students) have with the material.

During orientation week we verify that students have been able to load SAS and R onto their laptop computers.

The programming course begins with an orientation to SAS and R, which is intended to train students to perform basic tasks. It is important that students quickly learn to perform basic tasks in either language, because assignments in their first semester courses use both SAS and R.
Once the basics are completed, the course is divided into an R module and a SAS module. The R module focuses on writing (and using) R functions to perform analytical tasks such as simulations, graphics, and coding of statistical algorithms. The SAS module focuses on data management tasks such as restructuring data sets. Both modules provide practice in applying good programming principles (e.g., modularization) and are intended to provide instruction in how to think about programming in addition to the technical details of coding. Both modules also utilize principles of reproducible programming.

The programming sequence is new, and represents a significant upgrade from last year’s programming seminar, which was (among others) less structured.

Whether the programming sequence should be further upgraded – for example, to a 2- or 3-credit course – is an open question (and for which experience with the current course should provide guidance). Of particular importance is the need to provide SAS training at the level of SAS’s proficiency testing, as this would enhance the job prospects of our graduates. Developing a detailed competency test to be given at the conclusion of the course would assist in the assessment of competency.

The course sequence intends to combine “higher-level training” in software engineering (i.e., how to think about organizing and documenting a program) with the more mundane tasks of providing just-in-time instruction required to complete assignments from other courses (especially those from the applied data analysis sequence). The integration of these two goals will benefit from tighter collaboration among the course instructors.

Our tentative assessment is that the current course sequence would be adequate (entirely or substantially) for the first two years of doctoral study, but would then need to be supplemented by additional training in “computing”, which could include topics such as the technical details of how to perform analyses of large data sets and how to address precision in computationally-intense algorithms.
Outcome metrics:

- Alignment of the curriculum with desired program goals (by sequence)
  - Biology and communication: As two of the three core competencies in the program are biological knowledge and communication skill, the course content is integral to the curriculum.
  - Programming: This sequence is well-aligned with the overall program goals. It will be helpful to determine the degree to which our students would benefit from becoming formally credentialed in SAS, and whether any additional content would be required to obtain those credentials.
  - Probability and mathematical statistics sequence: This sequence provides the intellectual underpinnings of the applied data analysis sequence, and thus is well-aligned with the overall program goals.
  - Analytical methods: This sequence is directed toward the analytical competency, and thus is integral to the curriculum.

- Alignment of the curriculum with desired teaching methods (by sequence)
  - Biology and communication: This sequence is well-aligned with our advertised emphasis on active-learning-based education. It might be noted that these topics are particularly amenable to an active-learning based approach, and that the other core course sequences need not be taught in the same manner to be consistent with the program’s core educational principles.
  - Programming: This sequence is reasonably albeit not perfectly aligned with the tenets of active-based learning. A well-designed set of self-study assignments, perhaps designed to develop mastery in content areas that can’t be addressed during limited class time, would be consistent with those tenets.
  - Probability and mathematical statistics sequence: This sequence is reasonably aligned with the tenets of active-based learning – especially, their emphasis on learning from solving numerous (albeit standard) problems. The technical reports are a wonderful idea, which might be expanded. Class structure might be expanded beyond the traditional lecture format.
  - Analytical methods: This sequence is well-aligned with the tenets of active-based learning. Class structure might be expanded beyond the traditional lecture format.
• Alignment of the curriculum with desired student outcomes (by sequence)
  o Biology and communication: Although following more quantitative metrics will be helpful in the future, the student outcomes to date have been highly encouraging.
  o Programming: An evaluation of the current version of the course is premature. However, the current version does represent a reconceptualization of a previous seminar that did not produce student proficiency in data management at a level that the program desired.
  o Probability and mathematical statistics: The first-year student evaluations were particularly favorable in their assessment of the dedication and energy that the instructors put into their teaching. Most students appear to be learning as desired, but there are exceptions. The one student that dropped out of the program failed BIOS201. The one student that failed a portion of the qualifying exam failed the BIOS201/204 portion. Perhaps helpful could be an increased emphasis on early skills testing and remediation.
  o Analytical methods: The first-year student evaluations were quite positive. Students tended to do well on the qualifying exam. Students also noted that this sequence helped to prepare them for their internship.

Action items:

Biology and communication:

• Develop more quantitative assessment metrics.
• Revise and submit a manuscript describing the course (this will contribute toward the goal of becoming recognized as a leader in statistical education).

Programming:

• Develop a more detailed syllabus that includes a list of the programming techniques that we wish students to master.
• Compare this list against, for example, the content of the SAS credentialing tests.
• Develop competency tests for SAS and R.
Probability and mathematical statistics:
- More of the assignments might be upgraded to be more challenging – for example, to include elements of problem framing.
- Consider whether a TA could assist in grading.
- Experiment with alternatives to the traditional lecture format.

Applied data analysis:
- BIOS205 would benefit from a TA to assist in grading.
- Experiment with alternatives to the traditional lecture format.

General:
- Develop a map of the overall curriculum
## APPENDIX D

Curriculum mapping among competing programs, with reference to B&B proposed PhD program

<table>
<thead>
<tr>
<th>B&amp;B</th>
<th>UNC</th>
<th>HARVARD</th>
<th>U of WASHINGTON</th>
<th>JOHNS HOPKINS@</th>
<th>U of MICHIGAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intro Probability</td>
<td>Prob and Stat Inf I</td>
<td>Probability theory I</td>
<td>Stat Inference***</td>
<td>2-yr theory sequence</td>
<td>Prob and Dist Theory</td>
</tr>
<tr>
<td>Intro Biostat Methods</td>
<td>Intermed Stat Meth</td>
<td>Methods I</td>
<td>Biostatistics I</td>
<td>2-yr methods seq</td>
<td>Appl stat I: Lin Regr</td>
</tr>
<tr>
<td>Biostat Practice (bio and communication)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Biostatistical Investigations@ @</td>
</tr>
<tr>
<td>Intro stat computing</td>
<td>Intro stat comput</td>
<td></td>
<td></td>
<td></td>
<td>Statistical computing</td>
</tr>
<tr>
<td>Problems – seminar</td>
<td>Seminar (9 quarters)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intro Inference</td>
<td>Prob and Stat Inf II</td>
<td>Stat inference I</td>
<td>Stat Inference***</td>
<td>2-yr theory sequence</td>
<td>Biostat Inference</td>
</tr>
<tr>
<td>Intro Methods II</td>
<td>Intermed Lin Models</td>
<td>Methods II</td>
<td>Biostatistics II</td>
<td>2-yr methods seq</td>
<td></td>
</tr>
<tr>
<td>Biostat Practice II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stat computing II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems - seminar</td>
<td>Seminar (9 quarters)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Probability</td>
<td>Advanced probability and inference</td>
<td>Probability theory and applications II*</td>
<td>Advanced theory of statistical Inference**</td>
<td>2-yr theory sequence</td>
<td>Math stat I</td>
</tr>
<tr>
<td>Theory of Linear Models</td>
<td>Theory of Linear Models</td>
<td>Regression and Analysis of Variance*</td>
<td>Theory of Linear Models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Categorical data</td>
<td></td>
<td></td>
<td>Advanced methods Independent data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methods for learning and discovery (novel)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Problems – seminar</td>
<td>Seminar (9 quarters)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Inference</td>
<td>Advanced probability and inference II*</td>
<td>Statistical Inference II*</td>
<td>Advanced theory of statistical inference**</td>
<td>Math Stat II</td>
<td></td>
</tr>
<tr>
<td>Survival analysis</td>
<td>Survival analysis</td>
<td>Failure time data*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlated and Longitudinal</td>
<td>Longitudinal Data Analysis</td>
<td>Multivariate and longitudinal*</td>
<td>Advanced methods dependent data</td>
<td>ANOVA and linear mixed models</td>
<td></td>
</tr>
<tr>
<td>Problems – seminar</td>
<td>Seminar (9 quarters)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Inference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Harvard requires 4 of 6 courses, 5 of which are analogous to those included in our core requirements

** U of Washington is on quarter system and requires 3 quarters of this named course; **Taught by U of Washington Stat Dept

@ Hopkins is on quarter system and courses have generalized names that are difficult to map

@@ The one course among the five programs that is most similar to our Practice of Biostatistics; it is more concerned with statistical problems than biology
### Core requirements not among B&B core requirements

<table>
<thead>
<tr>
<th>B&amp;B</th>
<th>UNC</th>
<th>HARVAR</th>
<th>U of WASHINGTON</th>
<th>JOHNS HOPKINS</th>
<th>U of MICHIGAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biostatistics courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B&amp;B Elective</td>
<td>Generalized linear models</td>
<td>Regression modeling project</td>
<td>Generalized linear models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 credits upper level Bios courses</td>
<td>6 additional credits Bios courses</td>
<td>15 additional credits Bios courses</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stat consulting</td>
<td>Consulting seminar</td>
<td>Stat consulting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stat teaching</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Longitudinal Data Analysis</td>
<td></td>
<td></td>
<td></td>
<td>Advanced Calculus</td>
<td></td>
</tr>
<tr>
<td>Applied Survival</td>
<td></td>
<td></td>
<td></td>
<td>Stochastic processes</td>
<td></td>
</tr>
<tr>
<td>Other courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Principles of Epi for Public Health or Fundamentals of Epi</td>
<td>Epidemiologic Methods I or Fundamentals of Epi</td>
<td>9 credits selected from epidemiology list</td>
<td>Principles of Epidemiology</td>
<td>Strategies and Uses of Epidemiology or Principles and Methods of Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Principles of Epi for Public Health</td>
<td>Epidemiologic Methods</td>
<td>9 credits from epidemiology list</td>
<td>Principles of Epidemiology</td>
<td>Strategies and Uses of Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Principles of Epi for Public Health</td>
<td>Epidemiologic Methods</td>
<td>9 credits from epidemiology list</td>
<td>Principles of Epidemiology</td>
<td>Strategies and Uses of Epidemiology</td>
<td></td>
</tr>
<tr>
<td>Intro to biomedical sciences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Public Health Perspectives on Research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Additional core requirements

<table>
<thead>
<tr>
<th>B&amp;B</th>
<th>UNC</th>
<th>HARVARD</th>
<th>U of WASHINGTON</th>
<th>JOHNS HOPKINS</th>
<th>U of MICHIGAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 credits supporting program</td>
<td>6 credits supporting program</td>
<td>10 credits in a cognate field</td>
<td>18 credits outside the primary department</td>
<td>9 hours in a cognate area</td>
<td></td>
</tr>
<tr>
<td>TA one course per year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All programs require a course in responsible conduct of research
## APPENDIX E

### Biostatistics Graduate Program

<table>
<thead>
<tr>
<th>Description</th>
<th>FY14 BUDGET</th>
<th>FY15</th>
<th>FY16</th>
<th>FY17</th>
<th>FY18</th>
<th>FY19</th>
<th>FY20</th>
<th>FY21</th>
</tr>
</thead>
<tbody>
<tr>
<td>M.B. Tuition and Fees</td>
<td>(1,435,540)</td>
<td>(1,396,270)</td>
<td>(1,458,730)</td>
<td>(1,487,814)</td>
<td>(1,517,480)</td>
<td>(1,547,740)</td>
<td>(1,578,830)</td>
<td>(1,610,312)</td>
</tr>
<tr>
<td>Biostat Scholarships - M.B.</td>
<td>273,000</td>
<td>220,000</td>
<td>180,000</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
</tr>
<tr>
<td>PhD Tuition and Stipends - PhD 2nd yrs</td>
<td>0</td>
<td>0</td>
<td>175,084</td>
<td>178,585</td>
<td>182,157</td>
<td>232,250</td>
<td>236,895</td>
<td>193,307</td>
</tr>
<tr>
<td>PhD Tuition and Stipends - PhD 3+ yrs</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>200,000</td>
<td>200,000</td>
<td>200,000</td>
<td>200,000</td>
<td>200,000</td>
</tr>
<tr>
<td>Operating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching &amp; course development effort</td>
<td>388,305</td>
<td>414,718</td>
<td>430,433</td>
<td>436,261</td>
<td>442,207</td>
<td>464,021</td>
<td>470,206</td>
<td>476,515</td>
</tr>
<tr>
<td>Director of Graduate Studies</td>
<td>62,809</td>
<td>53,810</td>
<td>54,887</td>
<td>55,984</td>
<td>57,104</td>
<td>58,246</td>
<td>59,411</td>
<td>60,599</td>
</tr>
<tr>
<td>Program Administration</td>
<td>60,161</td>
<td>66,897</td>
<td>67,961</td>
<td>69,046</td>
<td>70,153</td>
<td>71,282</td>
<td>72,434</td>
<td>73,609</td>
</tr>
<tr>
<td>Academic Retreat</td>
<td>0</td>
<td>3,000</td>
<td>4,000</td>
<td>5,000</td>
<td>6,000</td>
<td>7,000</td>
<td>7,000</td>
<td>7,000</td>
</tr>
<tr>
<td>Other Operating</td>
<td>131,275</td>
<td>131,000</td>
<td>131,000</td>
<td>131,000</td>
<td>131,000</td>
<td>131,000</td>
<td>131,000</td>
<td>133,262</td>
</tr>
<tr>
<td>G&amp;A</td>
<td>19,277</td>
<td>20,083</td>
<td>20,648</td>
<td>20,919</td>
<td>21,194</td>
<td>21,946</td>
<td>22,235</td>
<td>22,530</td>
</tr>
<tr>
<td>SOM Space:</td>
<td>92,923</td>
<td>94,781</td>
<td>96,777</td>
<td>98,611</td>
<td>100,583</td>
<td>102,595</td>
<td>104,646</td>
<td>106,739</td>
</tr>
<tr>
<td>Contribution to B&amp;B Basic Operations</td>
<td>31,010</td>
<td>75,667</td>
<td>128,936</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
<td>150,000</td>
</tr>
<tr>
<td>Expenditures</td>
<td>1,058,760</td>
<td>1,079,957</td>
<td>1,289,625</td>
<td>1,495,407</td>
<td>1,510,398</td>
<td>1,588,341</td>
<td>1,604,948</td>
<td>1,573,561</td>
</tr>
<tr>
<td>Operating Variance</td>
<td>(376,780)</td>
<td>(316,313)</td>
<td>(169,104)</td>
<td>(7,593)</td>
<td>(7,083)</td>
<td>40,601</td>
<td>26,118</td>
<td>(36,751)</td>
</tr>
</tbody>
</table>

### Biostatistics Education Program Reserve

<table>
<thead>
<tr>
<th>Description</th>
<th>4310052</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning Balance</td>
<td>(301,659)</td>
</tr>
<tr>
<td>Revenues</td>
<td>24,250</td>
</tr>
<tr>
<td>Expenses/(Credits)</td>
<td></td>
</tr>
<tr>
<td>Transfers In</td>
<td>3,750</td>
</tr>
<tr>
<td>Transfers Out</td>
<td></td>
</tr>
<tr>
<td>Operating Variance - /(+ )</td>
<td>(376,780)</td>
</tr>
<tr>
<td>G&amp;A</td>
<td>728</td>
</tr>
<tr>
<td>Repay CRTP</td>
<td></td>
</tr>
<tr>
<td>Ending Balance</td>
<td>(649,712)</td>
</tr>
</tbody>
</table>
### Graduate Faculty

<table>
<thead>
<tr>
<th>Faculty Member</th>
<th>Tenure status</th>
<th>Rank</th>
<th>Role in program</th>
<th>Number of students as Dissertation Advisor</th>
<th>Number of PhD Committees</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Khalidi, Hussein</td>
<td>NT</td>
<td>Assoc</td>
<td>T, A, C</td>
<td>7</td>
<td></td>
<td>Came from industry, but served on Dissertation Committees at University of Cincinnati College of Pharmacy</td>
</tr>
<tr>
<td>Allen, Andrew</td>
<td>Tenured</td>
<td>Assoc</td>
<td>T, A, C</td>
<td>2</td>
<td>1</td>
<td>Has also supervised two postdocs: Min He (Assistant professor, Duke Center for Human Genome Variation) Chuanhua Xing (Assistant professor, Boston University)</td>
</tr>
<tr>
<td>Anstrom, Kevin</td>
<td>NT</td>
<td>Assoc</td>
<td>T, A, C</td>
<td>4</td>
<td>8</td>
<td>Huiman was a tenured faculty member at Emory prior to coming to Duke</td>
</tr>
<tr>
<td>Barnhart, Huiman</td>
<td>Tenured</td>
<td>Full</td>
<td>A, C</td>
<td>4</td>
<td>8</td>
<td>Two postdocs: Adam Richardson and Darongsae Kwon; also supervises two scientific programmers working on his research</td>
</tr>
<tr>
<td>Chan, Cliburn</td>
<td>NT</td>
<td>Asst</td>
<td>T, A, C</td>
<td>1</td>
<td>5</td>
<td>Supervised one postdoc: Jian Huang, currently at UCC, Ireland</td>
</tr>
<tr>
<td>Choudhury, Kingshuk</td>
<td>NT</td>
<td>Assoc</td>
<td>A, C</td>
<td>4</td>
<td>2</td>
<td>Has written several statistics textbooks and supervised several post-docs</td>
</tr>
<tr>
<td>Chow, Shein-Chung</td>
<td>NT</td>
<td>Full</td>
<td>A, C</td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>DeLong, Elizabeth</td>
<td>Tenured</td>
<td>Full</td>
<td>A, C – Director of Admissions</td>
<td>1</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Engelhardt, Barbara</td>
<td>TT</td>
<td>Asst</td>
<td>A, C</td>
<td>3</td>
<td>5</td>
<td>Also directing 2 post-docs</td>
</tr>
<tr>
<td>Erkanli, Alaattin</td>
<td>NT</td>
<td>Assoc</td>
<td>A, C</td>
<td>2</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Halabi, Susan</td>
<td>Tenured</td>
<td>Full</td>
<td>A, C</td>
<td>4</td>
<td>1</td>
<td>Member of Center for Human Genetics</td>
</tr>
<tr>
<td>Hauser, Elizabeth</td>
<td>NT</td>
<td>Full</td>
<td>A, C</td>
<td>4</td>
<td>9</td>
<td>Mentored a postdoc investigator, Insuk Sohn, for 3 years. He is working as a bioinformatician at Samsung Medical Center in Korea.</td>
</tr>
<tr>
<td>Jung, Sin-Ho</td>
<td>Tenured</td>
<td>Full</td>
<td>A, C</td>
<td>4</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Faculty Member</td>
<td>Tenure status</td>
<td>Rank</td>
<td>Role in program</td>
<td>Number of students as Dissertation Advisor</td>
<td>Number of PhD Committees</td>
<td>Notes</td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------</td>
<td>------</td>
<td>-----------------</td>
<td>------------------------------------------</td>
<td>--------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Koch, Allison</td>
<td>NT/secondary</td>
<td>Full</td>
<td>A,C</td>
<td>2</td>
<td>3</td>
<td>Member of Center for Human Genetics</td>
</tr>
<tr>
<td>Kosinski, Andrzej</td>
<td>TT</td>
<td>Assoc</td>
<td>T, A, C</td>
<td>2</td>
<td>7</td>
<td>Andrzej was a tenured faculty member at Emory prior to coming to Duke</td>
</tr>
<tr>
<td>Kuchibhatla, Maggie</td>
<td>NT</td>
<td>Assoc</td>
<td>T,C</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Li, Yi-Ju</td>
<td>NT</td>
<td>Assoc</td>
<td>A, C</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Li, Zhiguo</td>
<td>NT</td>
<td>Asst</td>
<td>T,A,C</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lokhnygina, Yuliya</td>
<td>TT</td>
<td>Asst</td>
<td>T,A,C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moser, Barry</td>
<td>NT</td>
<td>Full</td>
<td>T,A,C</td>
<td>9</td>
<td></td>
<td>I do not have a record of all of the graduate committees I served on as an advisor at Oklahoma State University. My estimate is that I served on one committee in the Statistics Department per year and one committee every other year in the Industrial Engineering Department for a period of 17 years.</td>
</tr>
<tr>
<td>Neelon, Brian</td>
<td>NT</td>
<td>Asst</td>
<td>A, C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neely, Megan</td>
<td>NT</td>
<td>Asst</td>
<td>T, C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O’Brien, Sean</td>
<td>TT</td>
<td>Asst</td>
<td>A, C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olsen, Maren</td>
<td>NT</td>
<td>Assoc</td>
<td>C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owzar, Kouros</td>
<td>NT</td>
<td>Assoc</td>
<td>T,A,C</td>
<td>3</td>
<td></td>
<td>Supervisor of two post-docs: Faheem Mitha Ivo Shterev</td>
</tr>
<tr>
<td>Pang, Herbert</td>
<td>TT</td>
<td>Asst</td>
<td>T, A, C</td>
<td>2</td>
<td></td>
<td>Senior Research Associate: Andrew Dellinger, PhD; Post-doc: Ace Hatch, PhD</td>
</tr>
<tr>
<td>Faculty Member</td>
<td>Tenure status</td>
<td>Rank</td>
<td>Role in program</td>
<td>Number of students as Dissertation Advisor</td>
<td>Number of PhD Committees</td>
<td>Notes</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
<td>------</td>
<td>-----------------</td>
<td>---------------------------------------------</td>
<td>--------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Pencina, Michael</td>
<td>Full</td>
<td></td>
<td>T, A, C</td>
<td>2</td>
<td>7</td>
<td>Michael Pencina has been appointed a Professor in the Department of Biostatistics and Bioinformatics at Duke University’s School of Medicine and Director of Biostatistics at Duke Clinical Research Institute, commencing in June 2013. Supervised 3 postdocs.</td>
</tr>
<tr>
<td>Peterson, Bercedis</td>
<td>NT</td>
<td>Assoc</td>
<td>T</td>
<td></td>
<td></td>
<td>See Notes</td>
</tr>
<tr>
<td>Pieper, Carl</td>
<td>NT</td>
<td>Asst</td>
<td>T, C</td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>Saha Chaudhuri, Paramita</td>
<td>NT</td>
<td>Asst</td>
<td>T, A, C</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Samsa, Greg</td>
<td>Tenured</td>
<td>Full</td>
<td>T, C - DGS</td>
<td></td>
<td></td>
<td>24</td>
</tr>
<tr>
<td>Turner, Liz</td>
<td>NT</td>
<td>Assistant</td>
<td>T, A, C</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Wang, Xiaofei</td>
<td>NT</td>
<td>Assoc</td>
<td>A, C</td>
<td>1</td>
<td></td>
<td>Postdoctoral advisor Junling Ma (Shanghai University of Finance)</td>
</tr>
<tr>
<td>Woolson, Robert</td>
<td>Adjunct</td>
<td>Full</td>
<td>A, C</td>
<td>18</td>
<td>&gt;30</td>
<td>Also directed a post-doctoral training program (T-32) at the University of Iowa for over 20 years and supervised multiple post-docs; currently working part-time at the VA and willing to take students</td>
</tr>
<tr>
<td>Wu, Yuan</td>
<td>NT</td>
<td>Asst</td>
<td>A, C</td>
<td>18</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Appendix G

Timeline

Upon entering the PhD program, each student will be assigned an initial advisor/mentor.

Typical First Year Schedule:

<table>
<thead>
<tr>
<th>Fall semester</th>
<th>Spring semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability (BIOSTAT 601)</td>
<td>Statistical Inference (BIOSTAT 604)</td>
</tr>
<tr>
<td>Biostatistical methods I (BIOSTAT 602)</td>
<td>Biostatistical methods II (BIOSTAT 605)</td>
</tr>
<tr>
<td>Biostatistical practice I (BIOSTAT 603)</td>
<td>Biostatistical practice II (BIOSTAT 606)</td>
</tr>
<tr>
<td>Intro to stat computing I (BIOSTAT 621)</td>
<td>Intro to stat computing II (BIOSTAT 622)</td>
</tr>
<tr>
<td>Problems in biostatistics - seminar</td>
<td>Problems in biostatistics - seminar</td>
</tr>
</tbody>
</table>

**First Year Examination:** At the end of the first year, all students take the methods examination. This exam covers topics spanning the range of core courses available during the first year and must be passed at a level sufficient to confirm the ability of the student in these core areas and, for the PhD student, their preparedness for progression to more advanced courses in the second year. As the material for the exam is taken from that covered in the first year courses, it is expected that students who successfully complete these courses will be well prepared for the exam.

**First Year Debrief:** At the end of the second semester and after the results from the methods exam are available, the student and their advisor will meet with the DGS to discuss the student’s progress, experiences, and interests going forward in the program. This will be an opportunity for the student to both give and receive feedback. Strategies for dealing with areas of weakness will be discussed. The interests and future plans of the student will be explored. If these interests are not consistent with the expertise of their assigned advisor, a new advisor will be assigned. This discussion will also be used to gauge the student’s likely course of study and will be used to inform the planning of future elective course offerings.

**First Year Summer Research Experience:** Students will be required to affiliate with program faculty for a directed research experience beginning the summer of their first year. The student’s advisor and the DGS will assist in this process. Having participated in the Problems in Biostatistics seminar during the first year, the students will have been exposed to a range of problem areas, as well as the faculty members working in them, that can provide dissertation topics.
Typical Second Year Schedule:

<table>
<thead>
<tr>
<th>Fall semester</th>
<th>Spring semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced Probability</td>
<td>Advanced Statistical Inference</td>
</tr>
<tr>
<td>Theory of linear models</td>
<td>Generalized linear models (BIOSTAT 719)</td>
</tr>
<tr>
<td>Categorical data analysis (BIOSTAT 714)</td>
<td>Survival analysis (BIOSTAT 713)</td>
</tr>
<tr>
<td>Statistical methods for learning and discovery (BIOSTAT 707)</td>
<td>Analysis of correlated and longitudinal data (BIOSTAT 718)</td>
</tr>
<tr>
<td>Problems in biostatistics -seminar</td>
<td>Problems in biostatistics - seminar</td>
</tr>
</tbody>
</table>

**Second Year Research Experience:** During the second year, students will complete two additional faculty directed research experiences (approximately one per semester).

**Second Year PhD Theory Examination:** At the end of the second year, PhD students take the theory examination. This exam covers topics spanning drawn from the second year core courses available and must be passed at a level sufficient to confirm the ability of the student’s preparedness for conducting independent methodological research. Again, as the material for the exam is taken from that covered in the second year core, it is expected that students that successfully complete these courses will be well prepared for the exam.

**Second Year Debrief:** At the end of the second year and after the results from the theory exam are available, the student and their advisor will meet with the DGS to discuss the student’s progress. The interests and future plans of the student will be explored including the likely direction of their research program and courses or experiences needed to round out their education.

**Third Year PhD Program (and beyond):** During the third year PhD students begin to specialize, acquiring the tools and techniques required for their chosen area of expertise. The student, with counsel from the DGS, will also select a primary dissertation advisor (if not selected earlier) and a dissertation committee. One member of the student’s committee will be either a clinical or a basic science investigator with expertise in subject matter relevant to the dissertation. With an advisor selected and a dissertation topic identified, the student and their advisor (with oversight/approval from the DGS) in collaboration with the DGS of the cognate program, will design and implement a cognate program designed to give the student additional background in a substantive area relevant to the students research topic. Since the program is tailored to the needs and career goals of each student, the exact program will differ for each student.
Typical Third Year Schedule (required courses in bold):

<table>
<thead>
<tr>
<th>Fall semester</th>
<th>Spring semester</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advanced topics in inference</td>
<td>Elective</td>
</tr>
<tr>
<td>Elective</td>
<td>Elective</td>
</tr>
<tr>
<td>Interdisciplinary elective</td>
<td>Interdisciplinary elective</td>
</tr>
<tr>
<td>Problems in biostatistics</td>
<td>Problems in biostatistics</td>
</tr>
</tbody>
</table>

Preliminary Examination: During the third year, PhD candidates will submit a short written manuscript on the research area to be presented and discussed at the exam, and which will usually be in their anticipated area of thesis research. Successful completion of the Exam qualifies the student as a PhD candidate in Biostatistics.

Yearly Debrief: Each year the student will meet with the student’s committee to discuss progress, as described above, with increasing emphasis on the supporting program and the dissertation.

The final oral: The student will present his/her thesis work at a seminar that will serve as the final oral, after which the student’s committee will vote to pass the student or to require additional work. Because of the review time (up to two years) involved for some statistical journals, the student will not be required to have the thesis published prior to being approved. However, it is expected that the student will have at least one draft statistical manuscript ready for submission by this time. Additionally, students should have one or more collaborative publications published or in press.

VI.0 New Course Offerings

Current problems in biostatistics (to be developed). Advanced seminar on topics at the research frontiers in biostatistics. This course is comprised of readings of current biostatistical research and presentations by faculty and advanced students of current research in their area of specialization. Required by all PhD students. Course instructor: rotating

Probability and Measure Theory (Sta 711 or developed de novo). Introduction to probability spaces, the theory of measure and integration, random variables, and limit theorems. Distribution functions, densities, and characteristic functions; convergence of random variables and of their distributions; uniform integrability and the Lebesgue convergence theorems. Weak and strong laws of large numbers, central limit theorem. Prerequisite: elementary real analysis and elementary probability theory. Prerequisites: Real Analysis. Instructor: TBD

Statistical Inference (Sta 732 or developed de novo) Classical, likelihood, and Bayesian approaches to statistical inference. Foundations of point and interval estimation, and properties of estimators (bias, consistency, efficiency, sufficiency, robustness). Testing: Type I and II errors, power, likelihood ratios; Bayes factors, posterior probabilities of hypotheses. The
predictivist perspective. Applications include estimation and testing in normal models; model choice and criticism. Prerequisites: STA 611 and STA 721 or equivalent. Instructor: Li, Wolpert

**Theory of Linear Models (Sta 721 or developed de novo).** Multiple linear regression and model building. Exploratory data analysis techniques, variable transformations and selection, parameter estimation and interpretation, prediction, Bayesian hierarchical models, Bayes factors and intrinsic Bayes factors for linear models, and Bayesian model averaging. The concepts of linear models from Bayesian and classical viewpoints. Topics in Markov chain Monte Carlo simulation introduced as required. Prerequisite: Statistics 213 and 290 or equivalent. Instructor: Clyde

**Advanced topics in modern inferential techniques and theory.** This course provides an introduction to topics that form the theoretical foundation of modern biostatistics and the inferential techniques for solving contemporary biostatistical problems. It covers topics in stochastic processes, weak convergence, semiparametric models and probability limit theory with an emphasis on those needed in biostatistics. Specific topics include stochastic processes, random walks, Markov chains, martingales, counting processes, Hilbert spaces for random vectors, semiparametric models, geometry of efficient score functions and efficient influence functions, the functional delta method, and the theory for M- and Z- estimators. Throughout, the presented theoretical results will be illustrated through application to real problems encountered during the faculty’s research. Additional topics including weak convergence and an overview of basic techniques in empirical process theory will be covered. This course provides students a ‘tool chest’ of advanced techniques for dealing with modern inferential problems. After completing this course, students will be well prepared for the inferential component of dissertation research. Prerequisites: Probability and measure theory (STAT 711) and statistical inference (STAT 732). Course instructor: Dr. Zhiguo Li
APPENDIX H

Curriculum Vitae of Faculty
Who Will Participate In the Program

(CVs start on next page)
NAME
Hussein R Al-Khalidi, PhD

POSITION TITLE
Associate Professor of Biostatistics

eRA COMMONS USER NAME (credential, e.g., agency login)
hushki

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mustansiriyah University, Baghdad, Iraq</td>
<td>B.S.</td>
<td>06/75</td>
<td>Statistics</td>
</tr>
<tr>
<td>Baghdad University, Baghdad, Iraq</td>
<td>M.S.</td>
<td>08/77</td>
<td>Statistics</td>
</tr>
<tr>
<td>Texas A&amp;M University, College Station, Texas</td>
<td>Ph.D.</td>
<td>08/88</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

Over the past 25 years, I have been actively conducting research in area of innovative statistical methodology, teaching graduate courses in statistics, serving on graduate thesis committees, developing curriculums, mentoring and serving on departmental committees. For the past 20 years, my main research focus has been on statistical methodology applications to medical sciences and especially to large cardiovascular clinical trials. In addition, I have 17 years of experience in applying innovated statistical methods to clinical trials with variety of composite and recurrent events endpoints. Over the years, I have been reviewer for several statistical and medical professional journals such as Communications in Statistics, Journal of the American Statistical Association, Journal of the American College of Cardiology, Journal of Biopharmaceutical Statistics, and Circulation.

B. Positions and Honors

Positions and Employment

1978-1982 Assistant Lecturer, Department of Statistics, Baghdad University, Baghdad, Iraq
1982-1985 Research Assistant, Department of Statistics, Texas A&M University, College Station, TX
1986-1988 Lecturer, Department of Statistics, Texas A&M University, College Station, TX
1988-1990 Assistant Professor, Department of Statistics, Baghdad University, Baghdad, Iraq
1990-1991 Visiting Research Associate, University of Texas at Tyler, Tyler, TX
1991-1993 Assistant Professor, Department of Epidemiology and Biostatistics, University of Illinois at Chicago, Chicago, IL
1993- Adjunct Associate Professor, College of Pharmacy, University of Cincinnati, Cincinnati, OH
1993-1997 Statistician, Department of Biometrics and Statistical Sciences, Procter & Gamble Pharmaceuticals, Cincinnati, OH
1997-2001 Senior Statistician, Department of Biometrics and Statistical Sciences, Procter & Gamble Pharmaceuticals, Cincinnati, OH
2001-2008 Principal Statistician, Department of Biometrics and Statistical Sciences, Procter & Gamble Pharmaceuticals, Cincinnati, OH
2009-2010 Director of Biostatistics, StatKing Consulting Inc., Cincinnati, OH
2010- Associate Professor, Dept. of Biostatistics and Bioinformatics, Duke University, Durham, NC

Professional Service

2009-2010 Secretary/Treasurer, Cincinnati Chapter, American Statistical Association
C. Selected Peer-reviewed Publications

Allen, Andrew S.

Associate Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center

Allen123

Institution and Location | Degree | Year(s) | Field of Study
--- | --- | --- | ---
University of New Mexico, Albuquerque, NM | B.S. | 1992 | Mathematics
University of New Mexico, Albuquerque, NM | M.A. | 1995 | Mathematics
Emory University, Atlanta, GA | Ph.D. | 2001 | Biostatistics

**A. Personal Statement**

For this application I will bring expertise in the development and application of statistical methodology that focuses on mapping complex human disease genes. Recent methodological work has included: adjusting for population stratification in case-control studies, testing untyped markers in whole genome association studies, estimating copy-number polymorphisms from intensity data, admixture mapping, and developing haplotype sharing approaches that can deal with computational requirements presented by genome-wide data. The breadth of my experience developing, evaluating, and applying statistical genetic methods in human genetic studies makes me well suited for the role highlighted in this application.

**B. Positions and Honors**

**Positions and employment**

- 1996-1997  Senior Statistician, Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM.
- 1997-2001  Teaching Assistant, Department of Biostatistics, Emory University School of Public Health, Atlanta, GA.
- 1997-2001  Research Assistant, Department of Biostatistics, Emory University School of Public Health, Atlanta, GA.
- 2001-2007  Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC.
- 2007-present  Associate Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC.

**Honors (selected)**

- Graduate Fellowship-Emory University
- Best Student Paper-1999 Joint Statistical Meetings, ASA Biopharmaceutical Section

**C. Selected peer-reviewed publications**


BIOGRAPHICAL SKETCH

NAME
Kevin J. Anstrom, Ph.D.

POSITION TITLE
Associate Professor

Department of Biostatistics and Bioinformatics
Duke Clinical Research Institute

eRA COMMONS USER NAME (credential, e.g., agency login)
Anstro001

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cornell University, Ithaca, NY</td>
<td>B.S.</td>
<td>05/1992</td>
<td>Statistics and Biometry</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>M.S.</td>
<td>05/1994</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>North Carolina State University, Raleigh, NC</td>
<td>Ph.D.</td>
<td>08/2002</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

Since joining the Duke faculty in the Fall of 2002, my research has been divided into two broad areas: 1) statistical analysis of observational and randomized data to explore the relationships between treatments, and long-term clinical, quality-of-life, and economic outcomes and 2) the design and analysis of randomized clinical trials using appropriate and novel statistical methods applied to idiopathic pulmonary fibrosis and heart failure patient populations.

B. Positions and Honors

Positions and Employment
1993-1995 Research Assistant, School of Public Health, University of North Carolina, Chapel Hill, North Carolina
1995-1997 Research Consultant, Department of Neurological Surgery, University of Washington
1997-2002 Statistician, Outcomes Research and Assessment Group, Duke University Medical Center, Durham, North Carolina
2002-2011 Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina
2011-Present Associate Professor, Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina

C. Selected Peer-reviewed Publications (out of 100+ peer-reviewed manuscripts)


BIOGRAPHICAL SKETCH

NAME
Huiman X. Barnhart

POSITION TITLE
Professor of Biostatistics

eRA COMMONS USER NAME (credential, e.g., agency login)
Huiman

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>South China Normal University, P.R. China</td>
<td>B.S.</td>
<td>1983</td>
<td>Mathematics</td>
</tr>
<tr>
<td>Jinan University, P.R. China</td>
<td>M.S.</td>
<td>1986</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of Pittsburgh, Pittsburgh, PA</td>
<td>M.A.</td>
<td>1988</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of Pittsburgh, Pittsburgh, PA</td>
<td>Ph.D.</td>
<td>1992</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

As the Principal Investigator of the Drug Induced Liver INjury Data Coordinating Center (DILIN DCC), I am ultimately responsible for the smooth operation of the DCC to meet the goals in the specific aims of DILIN DCC. I oversee the processes for monitoring site data, database validations and quality control efforts. I lead the DCC Statistical team and provide thought leadership and statistical advice on design and conduct of pharmacogenetic and biomarker studies as well as studies with intervention strategies to prevent and treat DILI activities. I also provide support to the network sites and assist the other steering committee members in the design, selection, and implementation of the network protocols.

I am currently a full professor in the department of Biostatistics and Bioinformatics. My research interests include both statistical methodology and disease-specific collaborative clinical research. My statistical expertise areas include methods for assessing reliability/agreement of measurement method, evaluating performance of new medical diagnostic tests, and methods for clinical trials. My work in the area of reliability will be useful for development of a validated, computer-based instruments. I have played pivotal leadership roles in major NIH-funded networks and consortia, and have gained considerable Data Coordinating Center (DCC) experience. I have been actively engaged in simultaneous development of multiple protocols within the network or consortium. I have provided input on the design and implementation of case report forms and databases, performed sample size calculations and re-calculation, advocated and led the implementation of sound statistical designs, overseen interim and final analyses, actively participated in manuscript writing, provided overall study coordination and quality assurance working with various groups within DCRI that supported the network or consortium, and has been an active member in almost all of the committees within the network or consortium. As the sole statistical faculty providing support on quality assessment and quality control for the Duke cardiovascular imaging core lab, I led the effort to revamp the process on assessment and control of quality and reliability on echocardiography measurements.

B. Positions and Honors

Positions and Employment

1992-1998: Assistant Professor, Department of Biostatistics, The Rollins School of Public Health, Emory University, Atlanta, GA
1998-2003: Associate Professor, Department of Biostatistics, The Rollins School of Public Health, Emory University, Atlanta, GA
2003-2009: Associate Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC
2009-present: Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC
C. Selected Peer-reviewed Publications (Selected from 105 peer-reviewed publications)

Most relevant to the current application


Additional publications (in chronological order)

NAME: Chan, Cliburn Chi Wei  
POSITION TITLE: Assistant Professor  

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>National University of Singapore, Singapore</td>
<td>MBBS</td>
<td>1986-1991</td>
<td>Medicine</td>
</tr>
<tr>
<td>University College London, London, UK</td>
<td>MSc, PhD</td>
<td>1997-2002</td>
<td>Nonlinear Dynamics</td>
</tr>
<tr>
<td>Imperial College London, London, UK</td>
<td>Postdoctoral</td>
<td>2002-2004</td>
<td>Immunology</td>
</tr>
<tr>
<td>Duke University, Durham, NC, USA</td>
<td>Postdoctoral</td>
<td>2004-2005</td>
<td>Computational Biology</td>
</tr>
</tbody>
</table>

Employment  
1992 – 1997  Resident, Singapore General Hospital and Singapore Armed Forces  
1999 – 2001  Teaching Assistant, University College London and London School of Economics  
2001 – 2004  Postdoctoral Fellow, Hammersmith Hospital, Imperial College London  
2004 – 2005  Research Fellow, Statistics and Applied Mathematics Institute (SAMSI)  
2005 - 2010  Assistant Professor of Biostatistics and Bioinformatics, Duke University  
2010 - 2011  Director, Biostatistics and Computational Biology Core, Duke Center for Aids Research  
2011 - 2018  Course director, Introduction to the Practice of Biostatistics I & II  

Honors and Synergistic Activities  
1998 – 2001  University College London Graduate School Award  
1998 – 2001  Overseas Research Award  
2002 – 2003  A*STAR (Singapore) International Fellow  
2007 - 2008  Reviewer for NSF, NIH (for RFA-AI-09-040: Protection of Human Health by Immunology and Vaccines (U01, U19), RFP NIAID-DMD-NIAID2010091, ‘Respiratory Pathogens Research Centers (RPRC), (ZAI1-UKS-A-J1), ‘Centers for AIDS Research’ etc)  
2008 - 2009  Invited speaker, Bioinformatics Summit, DAIDS, NIH; CMI Techniques Standardization for Vaccine Responses Evaluation, Fondation Merieux, Annecy, France; FICCS5 (Flow Informatics and Computational Cytometry Society), Seattle; Mathematics of Immunology and Infectious Diseases, AMS sectional meeting, NCSU; Cancer Vaccine Consortium (CVC) annual meeting, Washington D; CIMT special session on immuno-monitoring, Mainz, Germany.  
2009 - 2010  Instructor, Summer School on Quantitative and Computational Immunology, Santa Fe, NM, Summer School on Quantitative and Computational Immunology, San Antonio TX.  
2010 - 2011  Participant, inaugural Duke Course in Scientific Management and Leadership, Durham NC  
2011 - 2012  Instructor, Summer School on Quantitative and Computational Immunology, San Antonio TX  

Publications  
BIOGRAPHICAL SKETCH

NAME
Kingshuk Roy Choudhury

POSITION TITLE
Associate Professor

eRA COMMONS USER NAME (credential, e.g., agency login)
KRCHOUDHURY

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Statistical Institute, Calcutta, India</td>
<td>B.Stat.(Hons)</td>
<td>06/91</td>
<td>Statistics</td>
</tr>
<tr>
<td>Indian Statistical Institute, Calcutta, India</td>
<td>M.Stat.</td>
<td>06/93</td>
<td>Statistics</td>
</tr>
<tr>
<td>University of Washington, Seattle, WA, USA</td>
<td>Ph.D.</td>
<td>08/98</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

B. Positions and Honors

Positions and Employment
1998 – 2007  Lecturer, Statistics Department, University College Cork (UCC), Ireland
2007 – 2008  Visiting Professor, Department of Statistics, Stanford University
2008 – 2010  Lecturer, Statistics Department, UCC, Ireland
2011 – Present  Associate Professor, Dept. of Radiology, Duke University, Durham, NC, US
2013 – Present  Associate Professor, Dept. of Biostatistics and Bioinformatics, Duke University,

Honors and Committee Work

1993  P.C. Mahalanobis Memorial Statistics Symposium Gold Medal
1993  Indian Statistical Institute Alumni Association Gold Medal
2004-2007  Member, Executive committee, Irish Statistical Association
2009-2011  Member, Mathematical Sciences Committee, Royal Irish Academy
2012-Present  Member, Metrology committee, Quantitative Imaging Biomarker Alliance, Radiological Society of North America
2013-Present  Member, Working Group on Optimization of Medical Imaging Systems and Techniques, American Association for Physicians in Medicine

C. Selected Peer Reviewed Publications (selected from 44 full length journal publications)

2) O’Sullivan, F., Roy Choudhury, K., 2001, An Analysis of the Role of Positivity and Mixture Model
3) Roy Choudhury, K., Crotty, S., 2007, Morphometric analysis for early detection of changes in cellular structure in a toxicological experiment, Statistics in Medicine, 26, 29, 5253-5266


NAME
Shein-Chung Chow

POSITION TITLE
Professor, Biostatistics and Bioinformatics

eRA COMMONS USER NAME

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Taiwan University, Taipei, Taiwan</td>
<td>B.S.</td>
<td>1978</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of Wisconsin, Madison, Wisconsin</td>
<td>Ph.D.</td>
<td>1985</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

Have more than 25 years experience in pharmaceutical/clinical research and development. Have been working with many clinicians by providing statistical support to clinical studies conducted at early and/or late phases of clinical development across various therapeutic areas. Have published over 230 methodology papers and over 22 books. Have supervised 5 PhDs and 8 MS students. Serve on several editorial boards of leading journals in the area of biostatistics.

B. Positions and Honors

B.1 Positions and Employment

1985 - 1987  Assistant Professor, Department of Mathematics, University of Toledo, Toledo, Ohio
1989 - 1989  Senior Statistician, Parke-Davis Pharmaceutical Division, Ann Arbor, Michigan
1989 - 1997  Director, Biostatistics and Data Management, Bristol-Myers Squibb Company, Plainsboro, New Jersey
1992 - 2004  Honorary Affiliate Professor, Department of Statistics, Temple University, Philadelphia, Pennsylvania
1996 - Date  Adjunct Professor, Department of Statistics, National Cheng-Kung University, Tainan, Taiwan
2002 - 2004  Vice President, Biostatistics and Data Management, Millennium Pharmaceuticals, Inc., Cambridge, Massachusetts
2004 - 2005  Distinguished Professor, National Health Research Institutes, Zhunan, Taiwan.
2013 – Date  Adjunct Professor, Department of Statistics, North Carolina State University, Raleigh, NC
2005 - Date  Professor, Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC

B.2 Selected Honors

1995  Fellow, American Statistical Association
1996  The DIA Outstanding Service Award
1998  Chapter Service Recognition Award, American Statistical Association
1999  Elected Member, International Statistical Institute (ISI)

B.3 Editorial Boards

2002 – Date  Editor-in-Chief, Journal of Biopharmaceutical Statistics
1997 - Date  Editor-in-Chief, Book series "Biostatistics: A Series of Reference & Textbooks"
C. Peer-reviewed Publications

C.1 Selected Books (from over 22 books)


C.2 Selected Publications in Peer-reviewed Journals (from over 230 papers)


D. PhD Students Supervised

Dr. Wenping Wang (Temple University, 1995) – Co-advisor under special permission of Dean at Temple University
Dr. Li Li (Temple University, 2004) – Co-advisor under special permission of Dean at Temple University
Dr. Lan-Yan Yang (National Cheng-Kung University, 2011) – Co-advisor under special permission of dean at National Cheng-Kung University, Tainan, Taiwan
Ms. Ying Lu (Beijing Technology University, 2011 – Date) – Co-advisor under special permission of Dean of the university
Ms. Aijing Zhang (North Carolina State University, 2012 – Date) – Co-advisor since I was appointed as adjunct professor since March 1, 2013.
NAME: Elizabeth Ray DeLong Ph.D.

POSITION TITLE
Professor of Biostatistics & Bioinformatics

eRA COMMONS USER NAME: DELON001

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>Ph.D.</td>
<td>05/1979</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of Maine, Onoro, ME</td>
<td>M.A.</td>
<td>05/1970</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of Maine, Onoro, ME</td>
<td>B.A.</td>
<td>05/1969</td>
<td>Mathematics</td>
</tr>
</tbody>
</table>

Personal Statement

As Chair of the Department of Biostatistics and Bioinformatics, Duke University Medical Center and Co-Director of the Cardiovascular Outcomes Research group in the Duke Clinical Research Institute (DCRI), I participate directly in the selection of curriculum, course instructors, and advisors for our graduate programs. My research interests are in the field of comparative effectiveness with regard to cardiovascular outcomes and quality-of-care, with emphasis on risk adjustment methodology, assessment of risk prediction models, and provider profiling. With more than 20 years of biostatistics, clinical research, and bioinformatics experience, my responsibilities have included administrative and data analytic functions, as well as statistical methods development. I have held government grants studying statistical issues in Validating Risk Prediction Models in Cardiology and also Features of Managed Care Affecting Quality for Cardiovascular Disease and am currently the Principal Investigator for the analysis center for the ASCERT GO grant, a unique collaboration between the American College of Cardiology Foundation (ACCF) and The Society for Thoracic Surgeons (STS). I have taught several Biostatistics courses in the Medical School, including statistics for medical students, survival analysis, and a capstone course for the Clinical Research Training Program (CRTP).

Professional Positions

2009 – Present Chair, Department of Biostatistics and Bioinformatics, DUMC, Durham, North Carolina
2008 – 2009 Interim Chair, Department of Biostatistics and Bioinformatics, DUMC, Durham, North Carolina
2005 – Present Professor, Department of Biostatistics and Bioinformatics, DUMC, Durham, North Carolina
1997 – 2005 Associate Professor, Department of Biostatistics and Bioinformatics, DUMC, Durham, North Carolina
1996 - 1998 Chief, Division of Biostatistics and Clinical Outcomes, Department of Anesthesiology, Duke University Medical Center, Durham, North Carolina.
1994 - Present Co-Director Outcomes Research and Assessment Group, DUMC, Durham, North Carolina
1991 - 1997 Assistant Professor, Division of Biometry and Medical Informatics, Department of Community and Family Medicine, DUMC, Durham, North Carolina
1991 - 1994 Director Ischemic Heart Disease PORT, DUMC, Durham, North Carolina
1987 - 1991 Director of Biostatistics, Quintiles, Inc., Chapel Hill, North Carolina
1987 - 1991 Consulting Assistant Professor, Division of Biometry and Medical Informatics, Department of Community and Family Medicine, Duke University Medical School, Durham, North Carolina
1984 - 1987 Biostatistician, Health Services Research and Development Field Program, VA Medical Center, Durham, North Carolina
1979 - 1987 Assistant Professor, Division of Biometry and Medical Informatics, Department of Community and Family Medicine, Duke University Medical School, Durham, North Carolina
1979 - 1984 Biostatistician, Comprehensive Cancer Center, Duke University, Durham, North Carolina

Service Positions

2012 Grant Review Study Section, Ancillary Studies in Clinical Trials - National Heart, Lung and Blood Institute (NHBLI)
2012 The Joint Commission: Development of Risk Adjusted Outcome Measures in the EHR
C. SELECTED PEER – REVIEWED PUBLICATIONS (IN CHRONOLOGICAL ORDER).


BIOGRAPHICAL SKETCH

NAME
Barbara E Engelhardt

POSITION TITLE
Assistant Professor

eRA COMMONS USER NAME (credential, e.g., agency login)
ENGELHARDT

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stanford University, USA</td>
<td>BS</td>
<td>06/99</td>
<td>Symbolic Systems</td>
</tr>
<tr>
<td>Stanford University, USA</td>
<td>MS</td>
<td>06/99</td>
<td>Computer Science</td>
</tr>
<tr>
<td>University of California, Berkeley, USA</td>
<td>PhD</td>
<td>12/07</td>
<td>Computer Science</td>
</tr>
<tr>
<td>University of Chicago, USA</td>
<td>Postdoctoral</td>
<td>8/12</td>
<td>Human Genetics</td>
</tr>
</tbody>
</table>

A. Personal Statement

With a background in machine learning and statistics as applied to large biological data sets, I am uniquely qualified to develop and validate the statistical models required of the proposed research. I will develop the statistical methodology and analyze the results in a biological framework. I have experience with developing statistical models and software for genomics applications, including the development of the software SIFTER (supervised by Drs Michael I Jordan and Stephen E Brenner) to predict protein molecular function in protein families, and the development of sparse factor analysis models (supervised by Dr Matthew Stephens) to capture population structure using genotype data. More recently, I have completed work on several pharmacogenomic studies involving complex, high-dimensional phenotypes and gene expression data. In particular, I developed methodologies to perform association mapping with complex phenotypes associated with drug response, including models to identify differential expression quantitative trait loci (eQTLs). I am currently developing statistical models to build directed gene networks from gene expression data by incorporating available biological covariates (e.g., cis-eQTLs), and to identify individual-specific levels of mRNA isoforms from RNA-seq data in the independent phase of an R00. My new group at Duke University is focused on statistical methodologies for use in eQTL and association studies, joining forces with biologists to collect high-throughput biological and genomic data, and developing the appropriate statistical analysis to gain the most insight from jointly considering these diverse data.

B. Positions and Honors

Positions
1999-2001  Technical Staff, Artificial Intelligence Group, Jet Propulsion Laboratory, NASA
2001-2007  Graduate Student, Department of Computer Science, University of California, Berkeley
2007-2008  Scientist, 23andMe, Inc., Mountain View, California
2008-2011  Postdoctoral scholar, Department of Computer Science, University of Chicago
2011-2012  Postdoctoral scholar, Department of Human Genetics, University of Chicago
2012-present Assistant Professor, Department of Biostatistics & Bioinformatics, Duke University

Honors/Awards/Professional Activities
2001-2005  National Science Foundation Graduate Research Fellowship
2004    Walter M. Fitch Prize from Society for Molecular Biology and Evolution (SMBE)
2005-2006  Google Anita Borg Scholarship
C. Peer-reviewed publications


Date: March 12, 2013
Name: Alaattin Erkanli
Degree: Ph.D.
Gender: Male

1. **EDUCATION** (Begin with baccalaureate degree and include postdoctoral training, if applicable.)

<table>
<thead>
<tr>
<th>Institution and Location</th>
<th>Degree</th>
<th>Year Conferred</th>
<th>Scientific Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle East Technical University, Ankara, Turkey</td>
<td>B.S.</td>
<td>1981</td>
<td>Applied Statistics</td>
</tr>
<tr>
<td>Middle East Technical University, Ankara, Turkey</td>
<td>M.S.</td>
<td>1984</td>
<td>Mathematics</td>
</tr>
</tbody>
</table>

2. **PERSONAL STATEMENT**

I am a statistician who strives for finding sound methodological solutions to complex problems arising from real applications. I have over twenty years of on hand experience in providing statistical support to Duke Faculty and other collaborators. I routinely provide extensive power calculations for NIH/NIMH and similar grant applications using cutting edge statistical techniques. My research focus is in the applications of Bayesian and frequentist methodology to psychiatric epidemiology, mental health services, genetic epidemiology, multi-stage optimal designs, hierarchical generalized linear models, meta-analysis, structural equations (SEM), latent class models, item-response theory, non-parametric statistics, reliability, clinical cancer research, response adaptive randomization, and medical diagnostic testing. I made some important methodological contributions to statistical literature, and, through collaborative research with some of the best scientists in the world, I also made important contributions to psychiatric epidemiology literature.

2. **EXPERIENCE.**

1991-1993 Visiting Assistant Professor, Institute of Statistics and Decision Sciences, Duke University
1993-1994 Research Associate, Developmental Epidemiology Program, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine
1995-1999 Assistant Professor, Developmental Epidemiology Program, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine
1999-2006 Assistant Professor of Biostatistics, Department of Biostatistics and Bioinformatics, Duke University School of Medicine
2006- Associate Professor of Biostatistics, Department of Biostatistics and Bioinformatics, Duke University School of Medicine

3. **SELECTED PEER-REVIEWED PUBLICATIONS.** List (in chronological order) of representative publications

Erkanli, A. (1994). Laplace approximations for posterior expectations when the mode occurs on the
NAME
Susan Halabi, Ph.D.

POSITION TITLE
Professor with Tenure

eRA COMMONS USER NAME (credential, e.g., agency login)
Shalabi

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>American University of Beirut</td>
<td>B.S.</td>
<td>6/82</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>American University of Beirut</td>
<td>M.P.H.</td>
<td>6/84</td>
<td>Health Services Admin</td>
</tr>
<tr>
<td>The University of Texas Health Science Center at Houston, TX</td>
<td>Ph.D.</td>
<td>5/94</td>
<td>Biometry</td>
</tr>
</tbody>
</table>

A. Personal Statement

Since 1996, I have been actively engaged within the Geni-urinary Committee (GU) of the Alliance (formerly Cancer and Leukemia Study Group B). I have provided statistical support for investigators where I have been involved in study design, monitoring, study conduct, analyses, and manuscript writing. The collaboration was very productive and resulted in over several dozen publications in leading medical journals including the Journal of Clinical Oncology and Journal of the National Cancer Institute. In this application, I will collaborate with study chairs with the GU committee in the design, conduct and analysis of trials and correlative sciences studies. In addition, I have a tremendous interest and I maintain a research portfolio in prostate cancer, especially in developing and validating prognostic models. I have led multiple meta-analyses in prostate cancer and recently co-edited a book on clinical trials in oncology. Currently, I am a PI on a funded R01 award to develop prognostic models of clinical outcomes in men with castration resistant prostate cancer. I am also a PI on funded U01 award entitled ‘Validation of prognostic and pathway signatures in lethal prostate cancer. In summary, I have a demonstrated record of successful and productive research as a statistician and as an independent researcher within the GU committee.

B. Positions and Honors

List in chronological order previous positions, concluding with the present position. List any honors. Include present membership on any Federal Government public advisory committee.

7/94-12/96 Assistant Professor, Department of Biostatistics and Epidemiology, Tulane University
12/96-2/05 Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University
3/05-9/30/12 Associate Professor, Department of Biostatistics and Bioinformatics, Duke University
10/01/12-present Professor, Department of Biostatistics and Bioinformatics, Duke University
12/96-present Faculty Statistician, The Genitourinary Committee, Cancer and Leukemia Study Group B
6/03-06/06 Member, Central IRB (CIRB), NCI
6/03-6/06 Reviewer, Cancer Biomarker Study Section (CBSS), NCI
2007-2008 Ad hoc Member, Subcommittee H, NCI
08-present Member, GU Steering Committee, NCI
08-present Member, Prostate Cancer Task Force, NCI
2009-present Member, Editorial Board, Journal of Clinical Oncology
2009-present American Joint Committee on Cancer, Modeler Group
2008-2011 Member, Grant Selection Committee, American Society of Clinical Oncology
2011-2015 Member, Board of Directors, Society of Clinical Trials
2012-2013 Member and Director of Biostatistics Track (2013), American Society of Clinical Oncology Scientific Program Committee
11/11-present Associate Editor, Clinical Trials
C. Selected Peer-reviewed Publications (Selected from over 150 peer-reviewed publications)


NAME
Elizabeth R. Hauser, Ph.D.

POSITION TITLE
Professor

eRA COMMONS USER NAME (credential, e.g., agency login)
Elizabeth.Hauser

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ripon College, Ripon, WI</td>
<td>AB</td>
<td>05/1982</td>
<td>Biology</td>
</tr>
<tr>
<td>The Johns Hopkins University, Baltimore, MD</td>
<td>MHS</td>
<td>08/1985</td>
<td>Genet. Epidemiology</td>
</tr>
<tr>
<td>University of Michigan, Ann Arbor, MI</td>
<td>MS</td>
<td>05/1992</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of Michigan, Ann Arbor, MI</td>
<td>PhD</td>
<td>05/1998</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

I have a long-standing interest in understanding effects in cardiovascular disease and the relationship between cardiovascular disease, biomarkers for cardiovascular disease and genetic models for both cardiovascular disease and biomarkers. We have had tremendous success in the identification of genetic effects in the GENECARD study of early-onset coronary artery disease and we have recently turned our attention to understanding preclinical markers of cardiovascular risk in a cohort of individuals from the Duke Catheterization Lab. I am very excited about the emphasis of our ongoing studies to metabolomics biomarkers because this new direction allows for the potential for translation of our genetic and biomarker findings to treatments for coronary artery disease. My recent statistical methods development projects) in the setting of genetic heterogeneity (i.e. multiple genetic causes) has provided me an opportunity to explore the quantitative trait model, such as might be seen for genetic effects in metabolite profiles and the gene-by-covariate interaction models. Finally, I am familiar with the logistical aspects of performing a high-throughput genetic and genomics study. I direct the Duke Center for Human Genetics and supervise the Center for Human Genetics informatics group and as such I have been involved in the establishment of our WGA informatics and sequence analysis pipeline.

B. Positions and Honors

Positions and Employment

1985-1987 Research Assistant, The Health Services Research and Development Center, The Johns Hopkins University School of Hygiene and Public Health
1987-1988 Research Associate, Dept. Epidemiology and Preventive Medicine, University of Maryland at Baltimore School of Medicine
1988-1990 Research Associate, Dept. Pediatrics, The Johns Hopkins School of Medicine
1990-1997 Graduate Student Research Assistant, Dept. Biostatistics, University of Michigan School of Public Health
1998-2004 Assistant Research Professor, Section Medical Genetics, Dept. Medicine, Duke University Medical Center (DUMC), Durham, NC
2001-2004 Assistant Research Professor of Biostatistics and Bioinformatics, DUMC
2004-2010 Associate Professor, Division of Medical Genetics, Dept. Medicine, DUMC
2004-2010 Associate Research Professor of Biostatistics and Bioinformatics, DUMC
2007-present Adjunct Associate Professor, Dept. Statistics, College of Physical and Mathematical Sciences, NC State University, Raleigh, North Carolina
2007-present Associate Research Professor in Nursing, DUMC
2008-present Associate Research Professor in Statistical Science, Dept. of Statistical Science, Duke University, Durham, NC
2010-present Statistician, Durham VAMC, Durham, NC
2010-present Professor, Division of Medical Genetics, Dept. Medicine, DUMC
2010-present Research Professor of Biostatistics and Bioinformatics, DUMC
C. Selected peer-reviewed publications most relevant to this application (in chronological order from more than 85)


A. Personal Statement
I am the Director of Thomas Jefferson University’s Division of Biostatistics and the Director of the Biostatistics Core of Jefferson’s NCI-designated Kimmel Cancer Center. As of 1/2014, I will be the Director of Biostatistics for the Duke Cancer Institute, and Professor (pending final appointment) in the Department of Biostatistics and Bioinformatics at Duke. I have had more than 10 years of experience in cancer research, with a focus on estimation of and modeling of biomarkers, including mRNA, microRNAs and in situ tumor protein levels. I have NCI funding to support statistical investigations of quantification of biomarkers from q-RT-PCR. My current research focuses on sophisticated approaches to models of prognosis that utilize multiple biomarkers. I serve as the study biostatistician on several major collaborative cancer grants. For instance, a $ 6.7 million Komen Promise grant on breast cancer (quantitative in situ protein marker profiling and outcome, clinical trial) was the result of several years of weekly interactions with Dr. Hallgeir Rui (PI of Komen grant) and a multidisciplinary consortium of translational and clinician scientists.

B. Positions and Honors.
Positions and Employment
1991 - 2001 Promotion to Principal Biostatistician, Thomas Jefferson University, Department of Medicine, Philadelphia, PA
2001 - 2005 Assistant Professor, Thomas Jefferson University, Division of Biostatistics, Department of Pharmacology and Experimental Therapeutics
2003 - 2006 Assistant Director of Biostatistics Core, Thomas Jefferson University, Kimmel Cancer Center
2005-2013 Associate Professor, Thomas Jefferson University, Division of Biostatistics, Department of Pharmacology and Experimental Therapeutics
2006-2013 Director of Biostatistics Core, Thomas Jefferson University, Kimmel Cancer Center
2006-2013 Director, Division of Biostatistics, Thomas Jefferson University
2014-present Professor, Department of Biostatistics and Bioinformatics, Duke University
Director of Biostatistics, Duke Cancer Institute

Other Professional Experience
Reviewer, NCI Scientific Study Section (PARs-03-098, 099, SEP) 2003
Reviewer, NCI Parent Committee, Cancer Centers, Initial Review Group, ad-hoc 2007-2009
Reviewer, NCI SPORE (Specialized Programs of Research Excellence) 2007-present
applications in Breast, Gynecologic, Genitourinary, and Prostate Cancers 2007-present
Reviewer, NCI P01 Molecular Studies 2009-present
Reviewer, NCI Clinical Oncology (CONC) 2008-2009
Member, ASCO Scientific Program Committee 2009-present
Permanent Member, NCI, Subcommittee A, Cancer Centers 2009-2012
Reviewer, NCI Parent Committee, Cancer Centers, Initial Review Group, ad-hoc 2013-present
Statistical Adviser, Nature Publishing Group 2013-present
C. Selected peer-reviewed publications (in chronological order).
(Publications selected from >70 peer-reviewed publications)


DUKE UNIVERSITY MEDICAL CENTER
BIOSKETCH

Date Prepared: 3/11/2013

Name: Sin-Ho Jung, Ph.D.
Primary academic appointment: Professor (tenured)
Primary academic department: Department of Biostatistics and Bioinformatics

Education:

<table>
<thead>
<tr>
<th>Institution</th>
<th>Date [year]</th>
<th>Degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>College</td>
<td>1982</td>
<td>B.A.</td>
</tr>
<tr>
<td>Graduate or professional school</td>
<td>1984</td>
<td>M.S.</td>
</tr>
<tr>
<td>University of Wisconsin-Madison</td>
<td>1992</td>
<td>Ph.D.</td>
</tr>
</tbody>
</table>

Professional training and academic career: [chronologically, beginning with first postgraduate position]

1. Academic Appointments

<table>
<thead>
<tr>
<th>Institution</th>
<th>Position/Title</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo Medical/Graduate School</td>
<td>Assistant Professor</td>
<td>1994-1995</td>
</tr>
<tr>
<td>Hallym University</td>
<td>Assistant Professor</td>
<td>1995-3/99</td>
</tr>
<tr>
<td>University of Wisconsin-Madison</td>
<td>Visiting Assistant Professor</td>
<td>1/98-7/98</td>
</tr>
<tr>
<td>Indiana University (IU) School of Medicine</td>
<td>Associate Professor</td>
<td>8/98-9/01 (with tenure)</td>
</tr>
<tr>
<td>IU School of Public Health</td>
<td>Adjunct Faculty</td>
<td>8/99-9/00</td>
</tr>
<tr>
<td>Duke University</td>
<td>Associate Professor</td>
<td>10/01-5/31/06</td>
</tr>
<tr>
<td></td>
<td>Professor</td>
<td>6/1/06-present</td>
</tr>
</tbody>
</table>

2. Administrative Positions

<table>
<thead>
<tr>
<th>Institution</th>
<th>Position/Title</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>IU Cancer Center</td>
<td>Biostatistics Core Director</td>
<td>8/98-9/01</td>
</tr>
<tr>
<td>American College of Surgeons Oncology Group</td>
<td>Acting Group Statistician</td>
<td>4/02-4/03</td>
</tr>
<tr>
<td>Cancer &amp; Leukemia Group B</td>
<td>Director, Biostatistics Unit</td>
<td>1/08-present</td>
</tr>
<tr>
<td>Duke Cancer Institute</td>
<td>Interim Director, Biostatistics</td>
<td>4/11-present</td>
</tr>
</tbody>
</table>

3. Collaboration/Committee

<table>
<thead>
<tr>
<th>Institution</th>
<th>Position/Title</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mayo/North Central Cancer Treatment Group</td>
<td>Faculty Statistician</td>
<td>1992-1995</td>
</tr>
<tr>
<td>Hoosier Oncology Group</td>
<td>Faculty Statistician</td>
<td>1999-2001</td>
</tr>
<tr>
<td>IU Cancer Center, Scientific Review Committee</td>
<td></td>
<td>8/98-9/01</td>
</tr>
<tr>
<td>Mary Margaret Walther Program, Internal Advisory Committee</td>
<td></td>
<td>5/99-9/01</td>
</tr>
<tr>
<td>IU General Clinical Research Center, Advisory Committee</td>
<td></td>
<td>5/01-9/01</td>
</tr>
<tr>
<td>IU School of Medicine, IRB-02 and IRB-04</td>
<td></td>
<td>7/01-9/01</td>
</tr>
<tr>
<td>American College of Surgeons Oncology Group</td>
<td>Faculty Statistician</td>
<td>10/01-4/04</td>
</tr>
<tr>
<td>Cancer and Leukemia Group B</td>
<td>Faculty Statistician</td>
<td>5/04-present</td>
</tr>
<tr>
<td>North Central Cancer Treatment Group</td>
<td>Faculty Statistician</td>
<td>1/11-present</td>
</tr>
</tbody>
</table>

4. Professional Activities

<table>
<thead>
<tr>
<th>Institution</th>
<th>Position/Title</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>NICHD, NIH</td>
<td>Visiting Scientist</td>
<td>7/95-8/95</td>
</tr>
<tr>
<td>NSABP, Pittsburgh University</td>
<td>Research Scholar</td>
<td>1/96-1/96</td>
</tr>
<tr>
<td>NICHD, NIH</td>
<td>Visiting Scientist</td>
<td>7/96-8/96</td>
</tr>
<tr>
<td>NSABP, Pittsburgh University</td>
<td>Research Scholar</td>
<td>1/97-1/97</td>
</tr>
<tr>
<td>NCI Lymphoma Steering Committee, Member, 3/2010-present.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCI DLBCL Working Group, Member, 12/2010-present.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Publications:

Peer-Reviewed Papers (selected among N=139):

7. Jung SH. Randomized phase II trials with a prospective control, Statistics in Medicine, 2008; 27: 568-583

Chapters in books:

Books:
2. Jung SH. Randomized Phase II Clinical Trials in Oncology. CRC, expected to be in market in April 2013.

Editorials, position and background papers

Areas of research interest:
Statistical methods for cancer clinical trials, Survival analysis, Longitudinal data analysis, Design and analysis of phase II clinical trials, Bioinformatics, Clustered data analysis.
A. Personal Statement.

I am a genetic epidemiologist with over 15 years of experience in the genetic dissection of complex and Mendelian traits. I have a long-standing interest in the identification of gene*gene and gene*environment interactions and have applied a number of biostatistics and genetic analytic techniques to a host of human disease phenotypes, including neurologic, hematologic and ophthalmic. I have also been involved with methods development, particularly more recently to deal with the large dimensionality that is often present in genetic and genomic studies. With a wealth of applied projects, as well as methodologic projects, I can provide trainees with a rich professional perspective and internship possibilities. I have been involved in teaching at all levels (elementary and high school through postdoctoral and fellowships), through both didactic and one on one settings. I have a strong commitment to education and am also the Director of Graduate Studies for the Duke University Program in Genetics and Genomics since August of 2012. I have a secondary academic appointment in the Department of Biostatistics and Bioinformatics and am happy to be a part of such an important mission to train and engage future biostatisticians.

B. Positions and Honors.

Positions and Employment

1997-1998 ATPM Fellow, Office of Genetics and Disease Prevention, Centers for Disease Control and Prevention, Atlanta, GA.

1998-2000 Research Associate, Department of Medicine, Division of Neurology, Duke University Medical Center, Durham, NC.

2001-present Assistant Medical Professor, Department of Medicine, Section of Medical Genetics, Duke University Medical Center, Durham, NC.

2001-2007 Assistant Professor in Biostatistics and Bioinformatics, Section of Medical Genetics, Department of Medicine, Duke University Medical Center, Durham, NC.

2007-present Associate Professor, Department of Medicine, Section of Medical Genetics, Duke University Medical Center, Durham, NC.

Other Experience and Professional Memberships

1993-present Member, American Society of Human Genetics

2002-present Member, International Genetic Epidemiologic Society

2002 Reviewer, Mental Disorders Panel, Pennsylvania Department of Health

2006-2007 Reviewer, Special Emphasis Panel (ZMH1-ERB-S-06), NIH

2006 Reviewer, Italian Telethon Foundation

2007-present Member, American Society of Human Genetics Program Committee

2007-present Member, Chiari and Syringomyelia Foundation Scientific Board

2007 Reviewer, CPDD Study Section, NIH

2008 Reviewer, Mental Disorders Panel, Pennsylvania Department of Health

2009 Reviewer, Special Emphasis Panel (Z-DCI-SRB-R), NIH

2010 Reviewer, Study Section (ZES1-TN-J-R), NIH/NIEHS

2010 Reviewer, Special Emphasis Panel (ZDE1-VH [02]), NIH/NIDCR

Honors

1996 Sigma Xi Graduate Student Research Award, Emory University, Atlanta, GA
C. Selected peer-reviewed publications (in chronological order, out of over 95).


A. Personal Statement
While at the Emory University Rollins School of Public Health I was involved as the primary biostatistician on several major clinical trials. Since joining Duke University and the Duke Clinical Research Institute (DCRI) 9 years ago, I have become involved in the NIH sponsored Data Coordinating Centers for the PROMISE trial and the GUIDE-IT trial. I have been actively engaged in protocol development, overseen analyses, participated in the DSMB meetings, and collaborated in manuscript writing.

I have also been active in the comparative effectiveness research through the Duke Evidence-based Practice Center. Currently I am part of the Data Coordinating Center for the Drug Induced Liver Injury Network (DILIN) with focus on manuscript preparation, and I participate in analyses and manuscript preparation based on the Society of Thoracic Surgeons (STS) Database.

In addition to my collaborative research activities I have been active in statistical research. I have been principal investigator (PI) of statistical methodology research grants funded by the American Heart Association, a co-investigator on a NIH grant related to agreement methods, and PI of an internal Duke grant for statistical methodology research in the area of longitudinal analysis. I have advised several graduate students and have taught graduate biostatistics courses.

B. Positions and Honors

Positions and Employment
1990-1998  Assistant Professor, Department of Biostatistics, Emory University, Atlanta, GA
1999-2003  Associate Professor (tenure), Department of Biostatistics, Emory University, Atlanta, GA
2003-    Associate Professor, Dept. of Biostatistics & Bioinformatics, Duke University, Durham, NC

Other Experience and Professional Memberships
2005-2007  Member, Data Safety Monitoring Boards
1999-2012  Member, NIH Peer Review Committees: ad hoc reviewer

Honors
1983   Tuition scholarship from Vice Chancellors and Principals of the Universities of United Kingdom
1988   First prize at the student paper competition of 9th meeting of Society of Clinical Trials
C. Selected Peer-reviewed Publications (selected from 85 peer-reviewed publications)

Most relevant to the current application

5. Kosinski AS. A weighted generalized score statistic for comparison of predictive values of diagnostic tests. Statistics in Medicine, 2013 (in press; published online) DOI: 10.1002/sim.5587

Additional publications (in chronological order)

BIOGRAPHICAL SKETCH

NAME
Kuchibhatla, Maragatha

POSITION TITLE
Associate Professor

eRA COMMONS USER NAME (credential, e.g., agency login)
MAGGIEK

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>St. Francis College, Hyderabad, India</td>
<td>B.S.</td>
<td>06/79</td>
<td>Mathematics &amp; Physics</td>
</tr>
<tr>
<td>Osmania University, Hyderabad, India</td>
<td>M.S.</td>
<td>12/81</td>
<td>Statistics</td>
</tr>
<tr>
<td>Texas A &amp; M University, College Station, TX</td>
<td>Ph.D.</td>
<td>12/92</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement
As a faculty statistician, the majority of the studies I am involved in are with Department of Psychiatry, and the Duke Center for the Study of Aging. My primary statistical methodological interest, as evident from my publications, is longitudinal data analysis. I have acquired statistical expertise with regards to issues and modeling of cross-sectional and longitudinal analysis. I was a PI on a recently completed grant from the Duke Translational Medicine Institute on an application of a newer alternate statistical methodology on repeated measures data from a drug trial.

As a statistical investigator on several cycles of Conte Center for Neurosciences, a Center grant with the department of Psychiatry, and with the Duke Aging Center’s PEPPER Center grant, I am involved in a number of longitudinal analyses. I am also the primary statistician for a number of completed as well as ongoing drug and behavioral trials sponsored by both NIH and industry. The proposed study involves survival analyses and cross-sectional analysis of biomarkers. I have worked on similar studies and the analyses are straightforward.

B. Positions and Honors

Positions and Employment
1982-1987 Research Officer, Department of Statistics, National Institute of Rural Development, Hyderabad, India.
1987-1992 Graduate Assistant (Teaching and Research), Department of Statistics, Texas A & M University, College Station, Texas.
1993-1993 Statistician, Institute of Urban Affairs, North Carolina State University, Raleigh, NC.
1993-1996 Statistician, Center for Aging and Human Development, DUMC, Durham, NC.
1996-2008 Assistant Professor, Department of Biostatistics and Bioinformatics, DUMC, Durham, NC.
2008- Associate Professor, Department of Biostatistics and Bioinformatics, DUMC, Durham, NC.

Other Experience and Professional Memberships
1995 Member, American Statistical Association (ASA)
2000 Member, Gerontological Association of America (GSA)
2004 Affiliate, American Association of Geriatric Psychiatry (AAGP)
2009 Associate Editor - Journal of Nutrition for the Elderly

Honors
2008-2010 Grant reviewer – Alzheimer’s Association
2008 Duke Pepper Center’s Pilot Award on Statistical Methodology
2009 Biostatistical Methodology Grant from CTSA (Clinical and Translational Sciences Award)
C. Selected Peer-reviewed Publications


Additional recent publications of importance to the field (in chronological order)


BIOGRAPHICAL SKETCH

NAME

Yi-Ju Li

POSITION TITLE

ASSOCIATE PROFESSOR

eRA COMMONS USER NAME

YIJU.LI

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chung-Yuang Christian University, Taiwan</td>
<td>B.S.</td>
<td>06/1991</td>
<td>Mathematics</td>
</tr>
<tr>
<td>North Carolina State University, Raleigh, NC</td>
<td>M.S.</td>
<td>12/1994</td>
<td>Statistics &amp; Genetics</td>
</tr>
<tr>
<td>North Carolina State University, Raleigh, NC</td>
<td>Ph.D.</td>
<td>12/1996</td>
<td>Statistics &amp; Genetics</td>
</tr>
</tbody>
</table>

A. Personal Statement

I am a statistical geneticist with over 13 years of research experience in genetic studies of human complex diseases. My expertise includes genetic association method development, genetic data analysis, study design, and decision making for gene mapping research of complex diseases. I have devoted a significant part of my career in genetic studies of Alzheimer disease, Parkinson diseases, myopia, Fuchs Endothelial Cornea Dystrophy, and several clinical outcomes for patients underwent non-emergent coronary artery bypass grafting with cardiopulmonary bypass. In addition, my collaborators and I have developed a number of family-based association methods. I have also mentored graduate students by either serving as a thesis advisor or committee member, postdoctoral fellows, and a junior MD faculty by serving as a co-mentor for the K23 grant. As a faculty member in the Biostatistics and Bioinformatics Department, I am in strong support for the Ph.D. program and will be happy to mentor students from the proposed Ph.D. graduate program.

B. Positions and Honors

1996-1997 Postdoctoral Research Associate, Department of Dairy Science, Virginia Polytechnic Institute & State University, Blacksburg, Virginia

1998-2000 Research Associate, Department of Biosystems, The Graduate University for Advanced Studies, Hayama, Kanagawa, JAPAN

Oct 2000 - July 2002 Research Associate, Center for Human Genetics, Department of Medicine, Duke University Medical Center, Durham, North Carolina

Aug 2002 - Oct 2002 Instructor, Center for Human Genetics, Department of Medicine, Duke University Medical Center, Durham, North Carolina

Nov 2002 – Jan 2010 Assistant Professor, Track 5, Center for Human Genetics Department of Medicine, Duke University Medical Center, Durham, North Carolina

Mar 2007 – Jan 2010 Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina

Mar 2007 – present Adjunct Assistant Professor, Department of Statistics, North Carolina State University, Raleigh, North Carolina

Sept 2008 – Oct 2010 Adjunct Assistant Professor, Department of Community, Occupational & Family Medicine, National University of Singapore, Singapore

Feb 2010 – present Associate Professor, Track 5, Center for Human Genetics Department of Medicine, Duke University Medical Center, Durham, North Carolina

Feb 2010 – present Associate Professor, Track 5, Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina

Nov 2010 – present Adjunct Associate Professor, Duke-National University of Singapore Graduate Medical School, Singapore

Other Experience and Professional Services


2005 – 2006 Ad hoc reviewer for NINDS, NSD-K Study Section

2005 Reviewer for National Eye Institute, ZEY1 VSN Study Section

2005 Reviewer for Gene Variation & Evolution Study Section, Special Emphasis Panel ZRG1 GGG-H (60)
2007 Reviewer for GEI Study Investigators RFA 033, National Human Genome Research Institute Special Emphasis Panel
2007 – present External Reviewer for Research Grants Council of Hong Kong, China
2008 – 2012 Associate Editor, Molecular Biology and Evolution (MBE)
2011 – present Editorial Board Member, Journal of Biometrics & Biostatistics
2011 – present Editorial Board Member, ISRN Biomathematics
2013 Reviewer for 2013/05 ZRG1 BCMB-A (51) R for RFA RM11-006

C. Selected peer-reviewed publications (out of 76 publications)


126
BIOGRAPHICAL SKETCH

NAME
Zhiguo Li

POSITION TITLE
Assistant Professor of Biostatistics and Bioinformatics

eRA COMMONS USER NAME (credential, e.g., agency login)

EDUCATION/TRAINING

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xiangtan University, CHINA</td>
<td>B.S.</td>
<td>1993</td>
<td>Statistics</td>
</tr>
<tr>
<td>Wuhan University, CHINA</td>
<td>M.S.</td>
<td>1996</td>
<td>Statistics</td>
</tr>
<tr>
<td>University of Michigan, Ann Arbor, MI</td>
<td>PhD</td>
<td>2008</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of Michigan, Ann Arbor, MI</td>
<td>postdoc</td>
<td>2008-2010</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

My experience in clinical trials started when I was a postdoctoral research fellow at the University of Michigan, when I provided consultation to psychiatrists on the design of their clinical trials. During the same time, I also developed novel methodology for clinical trial design. As an assistant professor of Biostatistics and Bioinformatics at Duke University, I have been heavily engaged in cancer clinical trials especially clinical trials in hematological malignancies and statistical methods research related to cancer. With my extensive training and experience in clinical trials and statistical methodology, I am confident that I can achieve the goals of the proposed project.

B. Positions and Honors:

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Position/Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1989-1993</td>
<td>Department of Mathematics, Xiangtan University, CHINA</td>
</tr>
<tr>
<td>1993-1996</td>
<td>Department of Mathematics, Wuhan University, CHINA</td>
</tr>
<tr>
<td>1996-2002</td>
<td>Lecturer, Medical School of Peking University, CHINA</td>
</tr>
<tr>
<td>2002-2004</td>
<td>Department of Statistics, Florida State University, Tallahassee, FL</td>
</tr>
<tr>
<td>2004-2008</td>
<td>Department of Biostatistics, University of Michigan, Ann Arbor, MI</td>
</tr>
<tr>
<td>2008-2010</td>
<td>Postdoctoral Research Fellow, Institute for Social Research, University of Michigan, Ann Arbor, MI</td>
</tr>
<tr>
<td>2010-present</td>
<td>Assistant Professor, Dept. of Biostatistics and Bioinformatics, Duke University Medical Ctr., Durham, NC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Award/Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>Best First Year Student Award, Department of Statistics, Florida State University</td>
</tr>
<tr>
<td>2008</td>
<td>Distinguished Student Paper Award, ENAR Spring Meeting</td>
</tr>
</tbody>
</table>

C. Selected peer-reviewed publications (in chronological order):


BIOGRAPHICAL SKETCH

NAME  
Yuliya Lokhnygina

POSITION TITLE  
Assistant Professor of Biostatistics

eRA COMMONS USER NAME (credential, e.g., agency login)  
lokhn001

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moscow Institute of Physics &amp; Technology (MIPT), Moscow, Russia</td>
<td>BSc</td>
<td>1997</td>
<td>Control Theory &amp; Operations Research</td>
</tr>
<tr>
<td>MIPT, Moscow, Russia</td>
<td>MSc</td>
<td>1999</td>
<td>Control Theory &amp; Operations Research</td>
</tr>
<tr>
<td>Cornell University, Ithaca, NY</td>
<td>MSc</td>
<td>2001</td>
<td>Operations Research &amp; Industrial Engineering</td>
</tr>
<tr>
<td>North Carolina State University, Raleigh, NC</td>
<td>PhD</td>
<td>2004</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement
As a PhD-trained biostatistician with over 8 years of experience, I will bring the necessary statistical expertise to this project. I will help develop the appropriate statistical hypotheses to address clinical research questions, advise on the appropriate analytic methods, help interpret the results and contribute to reporting of the study findings in the manuscripts.

B. Positions and Honors

Positions and Employment
1999-2001  
Teaching Assistant, Dept of Operations Research and Industrial Engineering, Cornell University, New York

2001-2003  
Teaching Assistant, Dept of Statistics, North Carolina State University, North Carolina

05/2003-08/2003  
Intern, BioOncology Statistics group, Genentech, Inc.

09/2003-08/2004  
Intern, CT Statistics, Duke Clinical Research Institute, Durham, North Carolina

09/2004-present  
Assistant Professor of Biostatistics, Duke Clinical Research Institute and Department of Biostatistics and Bioinformatics, Duke University

09/2005-10/2006  
Biostatistician, Duke General Clinical Research Center

10/2006-present  
Biostatistician, Biostatistics Core, Duke Translational Medicine Institute

09/2010-present  
Associate Director, Biostatistics Core, Duke Translational Medicine Institute

Other Experience and Professional Memberships
American Statistical Association, International Biometric Society (ENAR)

Honors
Mu Sigma Rho
William Mendenhall Teaching Scholarship (2002)

C. Selected Peer-reviewed Publications  (from 36 peer-reviewed publications)


BIOGRAPHICAL SKETCH
Provide the following information for the key personnel in the order listed for Form Page 2. Follow the sample format for each person. DO NOT EXCEED FOUR PAGES.

NAME
Barry K. Moser, Ph.D.

POSITION TITLE
Associate Research Professor

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illinois Benedictine College, Chicago IL</td>
<td>B.S.</td>
<td>1973</td>
<td>Mathematics</td>
</tr>
<tr>
<td>Purdue University, W. Lafayette, IN</td>
<td>M.S.</td>
<td>1975</td>
<td>Statistics</td>
</tr>
<tr>
<td>Purdue University, W. Lafayette, IN</td>
<td>Ph.D.</td>
<td>1985</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement
My areas of expertise are in biostatistics theory and applications, particularly in the areas of cancer and transplant clinical trials research and global health initiatives. I have published a graduate level Statistical Linear Models textbook and published numerous statistical methodology papers in respected national journals. I have also published many collaborative research papers in areas such as: T cell diversity and T cell function in long-term human SCID chimeras, Partner-Assisted Emotional Disclosure for Patients with GI Cancer, Diagnosis of Genital Ulcer Disease in Malawi, and Racial Differences in Analgesic/Anti-inflammatory Medication Use. I am a faculty biostatistician with thirty-five years of methodological and collaborative experience. I am, or have been, the faculty statistician or the co-PI on various grants in several health related fields. I have been teaching Statistics at the graduate level since 1985 and I have been the major advisor for 9 PhD and 7 Masters’ degrees in Statistics.

B. Positions and Honors

Positions and Employment:
1985-1990 Assistant Professor, Department of Statistics, Oklahoma State University
1991-1997 Associate Professor, Department of Statistics, Oklahoma State University
1997-2002 Professor, Department of Statistics, Oklahoma State University
2002-present Associate Professor, Biostatistics and Bioinformatics, Duke University Medical Center

Organizations and Participation:
Member, American Statistical Association
Member, Society of Clinical Trials

Professional Awards and Special Recognitions:
Nominee for the AMOCO Foundation Award for Teaching Excellence, 1990
Nominee for the Burlington Northern Faculty Achievement Award, 1992
Nominee for the Graduate Student Association Outstanding Teaching Award, 1996
Graduate Student Association Outstanding Teaching Award, 1997
Oklahoma State University Celebrity for Outstanding Commitment to Teaching, 1997
Nominee for the Oklahoma Medal for Excellence in Teaching and Administration, 1998
C. Selected Peer-Reviewed Publications


**BIOGRAPHICAL SKETCH**

Provide the following information for the key personnel in the order listed on Form Page 2. Follow this format for each person. DO NOT EXCEED FOUR PAGES.

<table>
<thead>
<tr>
<th>NAME</th>
<th>POSITION TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neelon, Brian H</td>
<td>Assistant Professor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>eRA COMMONS USER NAME</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>bneelon</td>
<td></td>
</tr>
</tbody>
</table>

**EDUCATION/TRAINING**  
*(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)*

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/Y</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duke University, Durham, NC</td>
<td>BA</td>
<td>12/88</td>
<td>Philosophy</td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill</td>
<td>Teaching Certificate</td>
<td>06/93</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill</td>
<td>PhD</td>
<td>05/05</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>Harvard Medical School, Boston, MA</td>
<td>Post-Doctoral Fellowship</td>
<td>06/07</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

**A. Personal Statement**

As an Assistant Professor in the Department of Biostatistics and Bioinformatics at Duke University School of Medicine, Dr. Neelon divides his time between the Nicholas School of the Environment and the HSR&D Division at the Durham VA Medical Center. In this capacity, he provides statistical support for grant applications, conducts data analysis, and contributes to the development of new statistical methodology. His research focuses on Bayesian methods, longitudinal data analysis, cluster randomized trials, finite mixture models, zero-inflated models, and spatial statistics, with application to children’s health, chronic disease prevention, and health services research. He has published or co-authored numerous applied and methodological papers in these areas, and has served as Co-Investigator on several NIH-funded grants. Prior to his position at Duke, Dr. Neelon was a post-doctoral fellow in the Department of Health Care Policy at Harvard Medical School, where he developed new methods for analyzing health care utilization and medical expenditure data. Since arriving at Duke in 2009, he has collaborated closely with investigators from the Nicholas School and the Durham VA on several published manuscripts and recently funded grants.

**B. Positions and Honors**

**Positions and Employment**

<table>
<thead>
<tr>
<th>Year</th>
<th>Position and Employment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-2004</td>
<td>Graduate Research Assistant, Lineberger Comprehensive Cancer Research Center, University of North Carolina at Chapel Hill</td>
</tr>
<tr>
<td>2001-2003</td>
<td>Student Temporary Employee, Biostatistics Branch, National Institute of Environmental Health Sciences</td>
</tr>
<tr>
<td>2005-2007</td>
<td>Faculty Statistician, Center for Health Promotion and Disease Prevention, University of North Carolina at Chapel Hill (Joint Appointment with Department of Biostatistics)</td>
</tr>
<tr>
<td>2006-2007</td>
<td>Research Assistant Professor, Department of Biostatistics, University of North Carolina at Chapel Hill</td>
</tr>
<tr>
<td>2007-2009</td>
<td>Postdoctoral Fellow in Statistics, Department of Health Care Policy, Harvard Medical School</td>
</tr>
<tr>
<td>2009-2011</td>
<td>Associate in Research, Children’s Environmental Health Initiative, Nicholas School of the Environment, Duke University</td>
</tr>
<tr>
<td>2012-</td>
<td>Statistician, Durham VA Medical Center, Durham, North Carolina</td>
</tr>
<tr>
<td>2012-</td>
<td>Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University Medical Center</td>
</tr>
</tbody>
</table>

**Honors**

<table>
<thead>
<tr>
<th>Year</th>
<th>Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>Special Commendation: MS qualifying exams, Department of Biostatistics, University of North Carolina at Chapel Hill</td>
</tr>
<tr>
<td>1999</td>
<td>Delta Omega Society (National Public Health Honor Society)</td>
</tr>
<tr>
<td>2001-2003</td>
<td>NIEHS Student Temporary Employee Fellowship</td>
</tr>
</tbody>
</table>
2004  Kupper Dissertation Award (awarded by the University of North Carolina Department of Biostatistics for the best dissertation publication)

2012  Biometrics Showcase Selection, Joint Statistics Meeting, August 2012, San Diego, CA


C. Selected Peer-Reviewed Publications


BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2. Follow this format for each person. DO NOT EXCEED FOUR PAGES.

NAME
Megan L. Neely

POSITION TITLE
Assistant Professor

Department of Biostatistics and Bioinformatics
Duke University Medical Center

eRA COMMONS USER NAME (credential, e.g., agency login)

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clemson University, Clemson, SC</td>
<td>BS</td>
<td>2001-</td>
<td>Mathematical Sciences</td>
</tr>
<tr>
<td>Clemson University, Clemson, SC</td>
<td>MS</td>
<td>2005-</td>
<td>Mathematical Sciences</td>
</tr>
<tr>
<td>North Carolina State University, Raleigh, NC</td>
<td>PhD</td>
<td>2006-</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement
I currently teach two courses in the Masters program: BIOS 221 and 222, Introduction to Statistical Programming I and II. These courses are co-requisites for BIOS 202 and 205, respectively. In BIOS 221, I focus on programming in R. The course establishes a sound foundation in programming logic and the implementation of statistical simulations. In BIOS 222, I focus on programming in SAS. The course establishes good practices in data management and introduces computing approaches that reinforce reproducibility. Through my collaborative work at the Duke Clinical Research Institute, I have exposure to a wide variety of statistical methodological issues that arise when applying current approaches to data from both observational and clinical trial data (e.g. the need to tweak an existing method or develop a new approach). Thus, I think could serve as a co-chair or as a committee member for a PhD student.

B. Positions and Honors

Positions and Employment
2011 – Present Assistant Professor, Department of Biostatistics & Bioinformatics, Duke University, Durham, NC
2010 – 2011 Statistical Intern, Clinical Trials SIGMA Group, Duke Clinical Research Institute, Durham, NC
2007 – 2010 Adjunct Instructor, Department of Mathematics, Wake Technical Community College, Raleigh, NC
2004 – 2006 Research Assistant, Greenville Hospital University Medical, Greenville, SC

Other Experience and Professional Memberships
2011 – Present Statistical Editor, Journal of the American College of Cardiology Heart Failure
2009 – Present Member, Eastern North American Region of the International Biometric Society
2008 – Present Member, American Statistical Association

Honors
2009 Gertrude M. Cox Academic Achievement Award: Outstanding PhD Candidate, NC State University
2006 – 2010 NIH Fellowship: Biostatistics Training in the Omics Era, NC State University
2005 President’s Award: 4.0 Career GPA, Clemson University
2005 Summa cum Laude, Clemson University
2005 Martin Academic Achievement Award: Top Graduating Senior, Clemson University

C. Selected Peer-reviewed Publications and Abstracts


NAME
Sean M. O'Brien

POSITION TITLE
Assistant Professor of Biostatistics

eRA COMMONS USER NAME
OBRIE027

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haverford College, Haverford, PA</td>
<td>BA</td>
<td>1992</td>
<td>Religion</td>
</tr>
<tr>
<td>Southern College of Technology, Marietta, GA</td>
<td>MS</td>
<td>1995</td>
<td>Computer Science</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>PhD</td>
<td>1998</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A. Personal Statement
My career has focused on advancing statistical methodology and outcomes research within the Department of Biostatistics and Bioinformatics, the Duke Clinical Research Institute, and nationally. My scientific contributions are in the areas of health care performance evaluation, clinical risk prediction, and multi-center observational studies. I currently serve as the Statistical Director of the Society of Thoracic Surgeons (STS) Data Warehouse and Analysis Center at DCRI, a data warehouse that collects detailed clinical data from over 250,000 surgical patients annually from over 1000 sites. I am also Co-Principal Investigator of the Statistical and Data Coordinating Center for the NHLBI-funded ISCHEMIA Trial which has a planned enrollment of 8000 patients in 40 countries. Additionally, I have been an investigator on 4 federal grants and contracts that involved linking clinical registries with administrative data in order to study comparative effectiveness of cardiovascular interventions. My statistical expertise includes the analysis of multicenter observational data, development of risk prediction and propensity models, and Bayesian analysis. I also work on the development of quantitative methods for healthcare provider performance assessment, including the development of new composite measurement methodology used by Consumer Reports and the Society of Thoracic Surgeons for evaluating adult cardiac surgery providers. I have served on 12 thesis committees including 2 as chair. I am co-author of a textbook published by CRC press called “Exercises and Solutions in Biostatistical Theory” and of a forthcoming textbook called “Exercises and Solutions in Statistical Theory” to appear in 2013. My background includes a Ph.D. in biostatistics from the University of North Carolina at Chapel Hill and fellowship training in Bayesian methodology at the National Institute of Environmental Health Sciences.

B. Positions and Honors

Positions and Employment
1993-1994 English Teacher, Hong Ren High School and Joeshan Institute, Chiayi Taiwan
1995-2002 Graduate Research Assistant / Statistician II, UNC Lineberger Cancer Research Center, University of North Carolina, Chapel Hill
1996 & 1999 Teaching Intern, Department of Biostatistics, University of North Carolina, Chapel Hill
2002–2004 Research Fellow, National Institute of Environmental Health Sciences, Research Triangle Park, NC
2004–Pres Assistant Professor, Biostatistics and Bioinformatics, Duke University Medical Center, Durham NC

C. Selected Peer-reviewed Publications


BIOGRAPHICAL SKETCH

NAME
Maren K. Olsen, PhD

POSITION TITLE
Associate Research Professor

eRA COMMONS USER NAME
Olsen008

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Johns Hopkins University, Baltimore</td>
<td>B.S.</td>
<td>1994</td>
<td>Mathematical Sciences</td>
</tr>
<tr>
<td>Pennsylvania State University, University Park</td>
<td>M.A.</td>
<td>1996</td>
<td>Statistics</td>
</tr>
<tr>
<td>Pennsylvania State University, University Park</td>
<td>Ph.D.</td>
<td>1999</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement.
I have been the Director of the Durham VAMC Health Services Research Biostatistics Unit since June 2003. In this role, I supervise and mentor 5 masters-level statisticians and coordinate biostatistics support all of the research protocols in our center. My research activities have primarily focused on collaborative statistical methodology within the field of health services research. In these endeavors I have continued to apply and build expertise in three broad methodological areas: complicated longitudinal data structures (including clustered, longitudinal data), multiple imputation of missing data, and two-part models for longitudinal health data.

B. Positions and Honors.

Positions and Employment
2000-2010 Assistant Professor, Duke University Medical Center.
2000-present Biostatistician, Health Services Research & Development, VA Medical Center, Durham, NC
2003-present Director, Biostatistics Unit, Health Services Research & Development, VA Medical Center, Durham, NC
2010-present Associate Professor, Duke University Medical Center

Other Experience and Professional Memberships
2001-2002 Course Co-Director, Introduction to Statistical Methods, Clinical Research Training Program, Duke University Medical Center, Durham, NC
2003, 2005 Ad Hoc reviewer of the VA Scientific Review and Evaluation Board (SREB) Study Section
2009-2011 Course Co-Director, Clinical Research Seminar, Clinical Research Training Program, Duke University Medical Center, Durham, NC

Honors
1994 Mathematical Sciences Achievement Award, Johns Hopkins University

C. Selected peer-reviewed publications (in chronological order).

Most relevant to the current application
1. Olsen MK, DeLong ER, Oddone EZ, Bosworth HB. (2008). Strategies for analyzing multilevel cluster-
randomized studies with binary outcomes collected at varying intervals of time. *Statistics in Medicine*. 27, 6055-6071 PMID: 18825655


**Additional recent publications of importance to the field (in chronological order)**


BIOGRAPHICAL SKETCH

A. Personal Statement

As the lead statistician for the Pharmacology and Population Pharmacokinetics committee of the Alliance for Clinical Trials in Oncology (formerly the Cancer and Leukemia Group C; CALGB), I have over nine years of experience in designing and analyzing candidate marker and genome-wide pharmacogenomics studies in cancer. Four notable examples are GWAS studies CALGB 80303 (pancreas), 40101 (breast), 90401 (prostate) and 80405 (colon). I also have a proven record of designing and conducting clinical trials in oncology. Two notable studies are CALGB 100104, which assessed the effect of maintenance therapy in multiple myeloma (McCarthy et al; NEJM 2012), and ACOSOG Z9001 which assessed the effect of adjuvant imatinib in GIST (DeMatteo et al, Lancet 2009).

My main statistical research focus is the development of survival analysis methodology for genome-wide data. I am interested in diverse censoring mechanisms including singly and doubly interval, grouped and informative censoring. I conduct research in both the large and finite-sample size inference. This work started in the context of microarrays, later shifted to the GWAS setting and now has ventured in to the next generation sequencing realm.

I have been active in the research arena by developing open source technology Notable applications include from tools for conducting association studies using stream computing technologies (permPGU), tools for managing GWAS data (SNPpy) and tools for designing discovery and validation GWAS studies (survSNP).

As the director of the Bioinformatics Unit of the Alliance and the Biostatistics Core of the Radiation Countermeasures Centers of Research Excellence (RadCCORE), I have extensive experience in the administrative and scientific mentorship of research staff. Furthermore, as the lead instructor of the three courses offered at Duke University, I have extensive expertise in teaching.

In summary, I have a demonstrated record of successful and productive genomic research in the clinical trial setting. My leadership, technical and mentoring will stand me in good stead to lead the proposed research.

B. Positions and Honors:

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Position, Institution and Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993-1996</td>
<td>Department of Mathematics, Lehigh University. Bethlehem, PA</td>
</tr>
<tr>
<td>1996-1998</td>
<td>Department of Statistics, University of Chicago. Chicago, IL</td>
</tr>
<tr>
<td>1996</td>
<td>MET, Harrisburg Area Community College. Harrisburg, PA</td>
</tr>
<tr>
<td>1998</td>
<td>Department of Statistics, University of North Carolina. Chapel Hill, NC</td>
</tr>
<tr>
<td>1999</td>
<td>Clinical Pharmacology Data Sciences, Glaxo Wellcome. RTP, NC</td>
</tr>
<tr>
<td>2002-present</td>
<td>Assistant Professor, Dept. of Biostatistics and Bioinformatics, Duke University Medical Ctr., Durham, NC</td>
</tr>
<tr>
<td>2002-present</td>
<td>Faculty Statistician, The Cancer and Leukemia Group B (CALGB), Duke University Medical Center, Durham, NC</td>
</tr>
<tr>
<td>2002-2006</td>
<td>Faculty Statistician, ACOSOG, Duke University Medical Center, Durham, NC</td>
</tr>
<tr>
<td>2006-present</td>
<td>Director, Biostatistics and Computational Biology Core of RadCCORE</td>
</tr>
<tr>
<td>2006-present</td>
<td>Director, CALGB Bioinformatics Unit</td>
</tr>
</tbody>
</table>
C. Selected peer-reviewed publications (in chronological order):

**Most relevant to methods and software development for statistical genomics**


7. Sohn I, **Owzar K**, Lim J, George SL, Mackey Cushman S, Jung SH. Multiple testing for gene sets from microarray experiments. BMC Bioinformatics 2011:.209. PMCID: PMC3131260


**Most relevant to analysis of genomic data**


**Most relevant to the current application in terms of design and conduct of clinical trials**


NAME
Pang, Herbert

POSITION TITLE
Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University School of Medicine

eRA COMMONS USER NAME (credential, e.g., agency login)
oxbert

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yale University, New Haven, CT</td>
<td>Ph.D.</td>
<td>2008</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Positions and Honors.

Positions and Employment

9/03 – 12/03 Teaching Fellow, Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT
9/04 – 12/04 Teaching Fellow, Department of Epidemiology and Public Health, Yale University School of Medicine, New Haven, CT
8/08 – present Assistant Professor, Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC
8/08 – present Faculty Statistician, Alliance for Clinical Trials in Oncology (formerly Cancer & Leukemia Group B), Statistics and Data Center, Duke Cancer Institute, Durham, NC

Honors and Awards

2002 – 2006 Graduate Fellowship, Yale University
2006 Travel Award, 1st Annual NIH National Graduate Student Research Festival
2011 Burroughs Wellcome Fund Collaborative Research Travel Grant

Other Experience and Professional Memberships

2003 – present Member, American Statistical Association
2009 – present Member, American Association for Cancer Research
2010 – present Editorial Board, Journal of Biometrics and Biostatistics
2012 – present Editorial Board, Bioinformatics and Biometrics
2012 – present Editorial Board, Journal of Clinical Oncology

B. Selected peer-reviewed publications (in chronological order).


H. Pang, and H. Zhao (2008), Building Pathway Clusters from Random Forests Classification using Class Votes. BMC Bioinformatics 9:87. PMCID: PMC2335306

H. Pang, T. Tong, H. Zhao (2009), Shrinkage-based Diagonal Discriminant Analysis and its Applications in High-dimensional Data. Biometrics 65:1021-1029. PMCID: PMC2794982


BIOGRAPHICAL SKETCH

Provide the following information for the key personnel and other significant contributors. Follow this format for each person. **DO NOT EXCEED FOUR PAGES.**

<table>
<thead>
<tr>
<th>NAME</th>
<th>POSITION TITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michael J. Pencina</td>
<td>Professor of Biostatistics and Bioinformatics</td>
</tr>
<tr>
<td>eRA COMMONS USER NAME</td>
<td>pencina.michael</td>
</tr>
</tbody>
</table>

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing,)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Warsaw, Warsaw, Poland</td>
<td>M.A.</td>
<td>1998</td>
<td>Mathematics</td>
</tr>
<tr>
<td>Boston University, Boston, MA</td>
<td>Ph.D.</td>
<td>2003</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal Statement

The goal of this research is to develop and expand the existing methodology to address the problem of heterogeneity of treatment effects in randomized trials and registry studies. I have the expertise, leadership and motivation necessary to successfully carry out the proposed work. I have worked on topics related to risk assessment and treatment personalization in the field of cardiovascular disease prediction among participants enrolled in the Framingham Heart Study for the last 10 years. Construction of risk scores and evaluation of their performance is the area of my primary research interest. New methods proposed in my recent publications have been widely adopted as the most promising statistical tools in the assessment of model performance. Recent guidelines from the American Heart Association suggest reporting of results that utilizes metrics and presentation suggested in my work. My ongoing collaborations with thought leaders in the fields of cardiovascular risk prediction, imaging and assessment of usefulness of new markers further strengthen my ability to carry out this research. My publications record includes more than 150 manuscripts published in peer-reviewed journals. These experiences make me uniquely positioned to fulfill my role on this application.

B. Positions and Honors

**Positions and Employment**

<table>
<thead>
<tr>
<th>Years</th>
<th>Position</th>
<th>Institution and Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999 – 2003</td>
<td>Research Assistant, Teaching Assistant</td>
<td>Boston University, Statistics and Consulting Unit, Boston, MA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Department of Mathematics and Statistics, Boston, MA</td>
</tr>
<tr>
<td>2003 - 2008</td>
<td>Research Assistant Professor</td>
<td>Boston University, Department of Mathematics and Statistics,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Boston, MA</td>
</tr>
<tr>
<td>2007 - 2010</td>
<td>Senior Consulting Statistician</td>
<td>Harvard Clinical Research Institute, Boston, MA</td>
</tr>
<tr>
<td>2008 - 2013</td>
<td>Associate Professor</td>
<td>Boston University, Department of Biostatistics, Boston, MA</td>
</tr>
<tr>
<td>2011 - 2013</td>
<td>Director, Statistical Consulting</td>
<td>Harvard Clinical Research Institute, Boston, MA</td>
</tr>
<tr>
<td>2013 -</td>
<td>Professor</td>
<td>Duke University, Biostatistics and Bioinformatics, Durham, NA</td>
</tr>
<tr>
<td>2013 -</td>
<td>Director</td>
<td>Duke Clinical Research Institute, Biostatistics, Durham, NA</td>
</tr>
</tbody>
</table>

**Other Experience and Professional Memberships**

<table>
<thead>
<tr>
<th>Years</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012 -</td>
<td>Associate Editor, Statistics in Medicine</td>
</tr>
<tr>
<td>2007 - 2011</td>
<td>Assistant Editor, Statistics in Medicine</td>
</tr>
<tr>
<td>2005 -</td>
<td>Member, American Heart Association</td>
</tr>
</tbody>
</table>
C. Selected peer-reviewed publications (Selected from over 100 publications)

Most relevant to the current application


Additional recent publications of importance to the field (in chronological order)


NAME
Bercedis L. Peterson

POSITION TITLE
Associate Research Professor
eRA COMMONS USER NAME (credential, e.g., agency login)
b peter

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Southern Illinois University</td>
<td>B.A</td>
<td>12/73</td>
<td>English Literature</td>
</tr>
<tr>
<td>Southern Illinois University</td>
<td>M.S.</td>
<td>6/77</td>
<td>Measurement and Stats</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill</td>
<td>M.S.</td>
<td>8/82</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill</td>
<td>Ph.D.</td>
<td>8/86</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal Statement:

I am fully funded by grants within the Cancer Institute where I have worked since 1991. I collaborate with Cancer Institute members in the design of grants, clinical trials, retrospective research projects, biomarker studies, and laboratory studies. If a project is funded I serve as the faculty statistician on the grant. From 1991 to 2008 I worked on the Leukemia and Breast Committees of the Cancer and Leukemia Group B, designing, monitoring, analyzing, and publishing Phase I, II and III clinical trials for CALGB. While there is currently limited opportunity for me to work on Phase III clinical trials at Duke, I make up for this by working with members of the Cancer Prevention, Detection, and Control division of the Cancer Institute to design and run complex randomized intervention studies.

B. Positions and Honors:

2005-present  Associate Research Professor, Department of Biostatistics and Bioinformatics, DUMC
2000-2005  Assistant Research Professor, Department of Biostatistics and Bioinformatics, DUMC
1984-88  Biostatistician, Behavioral Medicine Research Center, DUMC
1983  Graduate Research Assistant, EPA, Clinical Studies, University of North Carolina
1983-82  Graduate Research Assistant, Institute of Research in Social Sciences; UNC
1977-79  Staff Researcher, Computing Affairs Research and Evaluation Center, Computer Center, Southern Illinois University

C. Selected Peer-reviewed Publications from over 100.


2. Peterson, BL, George, SL. Sample size requirements and length of study for testing interaction in a 2xk factorial design when time-to-failure is the outcome. *Controlled Clinical Trials* 14:511-211, 1993.


A. Personal Statement.

I have worked as an applied biostatistician in gerontology and aging for over 20 years – directing the Computing and Statistics Laboratory for the Center on Aging and Human Development at Duke, and as a faculty member in the Duke Clinical Research Institute, for that entire time. The research I have participated in spans from macro sociological studies to cellular. In that capacity, my lab and I have provided design and analytic consultation on a wide variety of studies dealing with issues in aging, including clinical trials, analysis of longitudinal epidemiologic studies, pilot studies, meta-analysis, and administration Center grants. I list over 250 publications in refereed journals, many of which deal with analysis of exercise, function and performance data, and issues of aging. I have directed analysis and data management cores for Duke’s Older American Independence Center (Pepper OAIC) for three 5 year cycles, including work on the analytic issues in the study of metabalomic and physiologic biomarkers as they relate to aging, physical performance, and function.

I have served as the core director for numerous other center grants, including PI of the Data Analysis and Management core for Duke's Bryan Alzheimers Disease Center, PI of the Analysis and Data Management core for Duke's department of Psychiatry's Conte Center for the study of depression in the elderly for 2 cycles, and I was also PI for a twenty five million dollar contract for NIDA for the data management center for the Institute's Clinical Trials Network, directing the data management activities for all clinical trials on drug abuse conducted within the network.

My substantive research interests include psychometric/data aggregation methodology specific to metabolomics, analysis of function and performance, and methodological issues in methods for repeated measures designs, methods for data reduction, and issues in the assessment of reliability. I have been the lead statistician in works on aging in a variety of NIH projects (clinical trials, cross sectional, longitudinal, observational, and meta-analyses of secondary data sources). In addition, I have mentored students on K-awards and post-doctoral work, mostly dealing with issues in aging, on PhD committees providing statistical guidance on issues in physical performance and other aging problems, and as a statistical advisor on over 40 master’s essays in the Clinical Research Training Program (CRTP) in Duke Medical Center. I am a expert in reliability, clinical trials, design and longitudinal data analysis – having published both methodologic and substantive pieces in these areas, as well as teaching courses in the CRTP program in both psychometrics and longitudinal data analysis as well as Observational Data Analysis in the Biostatistics and Bioinformatics masters program.

Positions and Honors

1984-87 Adjunct Instructor in Biostatistics, School of Health Sciences, Hunter College, City U. of NY
1984-89 Adjunct Lecturer, Division of Biostatistics, Columbia University School of Public Health.
1990 Adjunct Assistant Professor, Div. of Biostatistics, Columbia Univ. School of Public Health
1990-91 Assistant Professor of Statistics in Medicine, Dept. of Medicine, NY Hosp/Cornell Medical Ctr
1994-Pres. Senior Fellow, Duke University Center for Aging and Human Development
2001-Pres. Assistant Professor, Dept. of Biostatistics and Bioinformatics, Duke Univ. Med. Ctr
C. Selected Peer-reviewed Publications  (Selected from 250+ peer reviewed publications)

Most relevant to the current application


3. Pan, JJ, Nahm,M, Wakim,P, Cushing,C, Poole,L, Tai,B, **Pieper,C**, 'A centralized informatics infrastructure for the National Institute on Drug Abuse Clinical Trials Network' *Clinical Trials Journal of the Society for Clinical Trials*, 2009: 6(1), 67-75. PMCID: PMC2962616


Additional recent publications of importance to the field


15. Morey,MC, Peterson,MJ, **Pieper,CF** et al., The Veterans Learning to Improve Fitness and Function in Elders Study: A Randomized Trial of Primary Care Based Physical Activity Counseling For Older Men, *Journal of the American Geriatrics Association*, 2009: 57(7), 1166-1174. PMCID: PMC2757328
BIOGRAPHICAL SKETCH

NAME
Paramita Saha Chaudhuri

eRA COMMONS USER NAME
CHAUDHURI001

POSITION TITLE
Assistant Professor of Biostatistics

EDUCATION/TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presidency College, Kolkata, India</td>
<td>BSc</td>
<td>1999</td>
<td>Statistics</td>
</tr>
<tr>
<td>Indian Statistical Institute, Kolkata, India</td>
<td>MStat</td>
<td>2001</td>
<td>Statistics</td>
</tr>
<tr>
<td>University of Washington, Seattle, WA</td>
<td>MS</td>
<td>2005</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of Washington, Seattle, WA</td>
<td>PhD</td>
<td>2009</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal statement

My research interest includes risk prediction, assessment of prediction accuracy, efficient design for observational studies and gene by environment interaction in longitudinal setting. I am experienced in analysis of observational studies with epidemiological, correlated, and survival data, in particular registry data (cardiac surgery registry) and electronic health records (diabetes).

B. Positions and Honors

Positions and Employment

2001 – 2003  Research Fellow, Theoretical Statistics and Mathematics Unit, Indian Statistical Institute, Delhi, India.
2004 – 2009  Research Assistant, Teaching Assistant and Consultant, University of Washington, Seattle, WA.
2009 – 2011  Research Fellow, National Institute of Environmental Health Sciences, Research Triangle Park, NC.
2011 – Present  Assistant Professor, Biostatistics and Bioinformatics, Duke University Medical Center, Durham NC.
2011 – Present  Faculty Statistician, Outcomes Research and Assessment Group, Duke Clinical Research Institute, Durham, NC.

Other Activities

2004 – Present  Member, American Statistical Association
2006 – 2009  Member, Western North American Region, International Biometric Society
2009 – Present  Member, Eastern North American Region, International Biometric Society
2009 – Present  Member, Society for Epidemiologic Research
C. Selected Peer-reviewed Publications (chronological order)


NAME  
Samsa, Gregory  

POSITION TITLE  
Associate Professor  

eRA COMMONS USER NAME (credential, e.g., agency login)  
gsamsa  

EDUCATION/TRAINING  
(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)  

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>B.A.</td>
<td>1979</td>
<td>Mathematics</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>M.S.</td>
<td>1981</td>
<td>Statistics</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill, NC</td>
<td>Ph.D.</td>
<td>1988</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal Statement  
I am an applied statistician with interests in information synthesis, practice improvement and system design, randomized trials, outcomes research, and most particularly in improving the link between evidence and practice. I have published on both the theory and practice of quality improvement and practice improvement research, and has served as the lead statistician on numerous studies that have developed and assessed the impact of putative improvements to health care systems. I am also committed to mentorship – for example, in my role as Director of Graduate Studies in the Department of Biostatistics and Bioinformatics, and through my work with various junior investigators.

B. Positions and Honors  
1980-1988 Statistician, Department of Infectious Diseases, University of North Carolina  
1988-1995 Senior Research Scientist, Department of Veterans Affairs Medical Center, Durham NC  
1988-1998 Assistant Professor, Duke University Division of Biometry, Dept of Community and Family Medicine  
1993-2009 Associate Director, Duke University Center for Health Policy Research and Education  
1991-present Assistant Research Professor, Duke University Division of General and Internal Medicine, Dept of Medicine  
1998-present Associate Professor, Duke University Division of Biometry, Dept of Biostatistics and Bioinformatics  
2009-present Director of Graduate Studies, Duke University Division of Biometry, Dept of Biostatistics and Bioinformatics

C. Selected Peer-reviewed Publications (from approximately 200, 25 first-authored)  
BIOGRAPHICAL SKETCH

NAME
Turner, Elizabeth

POSITION TITLE
Medical Instructor, Duke Global Health Institute and Department of Biostatistics and Bioinformatics, Duke University, USA

eRA COMMONS USER NAME
ELTURNER16

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warwick University, UK</td>
<td>BSc</td>
<td>07/00</td>
<td>Mathematics</td>
</tr>
<tr>
<td>McGill University, Canada</td>
<td>MSc</td>
<td>10/02</td>
<td>Statistics</td>
</tr>
<tr>
<td>McGill University, Canada</td>
<td>PhD</td>
<td>10/07</td>
<td>Statistics</td>
</tr>
</tbody>
</table>

A. Personal statement:
The key experiences and abilities that I offer include a depth and breadth of analytical skills including cluster-randomized trials, longitudinal cohort data, mediation and moderator analyses of both trial and observational data, and a range of outcome measures including anthropometric, cognitive, educational, psychological and socioeconomic measures.

As research fellow in the Department of Medical Statistics at the LSHTM, my role as collaborative research statistician with six main research groups has led to the development and honing of the key skills outlined above. These include Preventive Cardiology, through Imperial College London, with a focus on the European cluster randomized trial of a family-based preventive cardiology programme (ISRCTN 71715857); Public Health Entomology with a focus on prevention of malaria transmission through innovative transmission prevention strategies in Tanzania; International Centre for Eye Health with a focus on disease burden estimation, particularly due to cataract blindness, which led to the development of new statistical methodology (manuscript in review); European Huntington's Disease Network with the goal to use longitudinal registry data to inform clinical trial design; Stroke Research, through University College London, with a focus on the International Carotid Stenting RCT. In recognition of my analytical and collaborative skills, I was invited to become statistician to the Health and Literacy Intervention study (HALI; NCT00878007), a cluster-randomized trial of 5100 children in 101 schools in Kenya. HALI aims to evaluate a school-based malaria treatment to improve educational achievement.

B. Positions and Honors

Positions and Employment

2003-2007  Research Assistant, Lady Davis Institute for Medical Research, Montreal
2007-2012  Research Fellow, Department of Medical Statistics, Faculty of Population Health, London School of Hygiene and Tropical Medicine, London, UK
2012-      Medical Instructor, Duke Global Health Institute and Department of Biostatistics and Bioinformatics, Duke University, US

Honors and awards

2000-2001  Entente Cordial Scholarship, France (declined)
2000-2001  Canada Memorial Foundation Scholarship
2001-2003  Institut des Sciences Mathématiques Scholarship, Canada
2003-2005 Commonwealth Scholarship, Canada
2003-2006 McGill Major Scholarship, Canada (declined first two years)
2005 One of top three presentations, Research Students Conference, Cambridge

Committees and Memberships
2007- Member, Royal Statistical Society, UK
2011- Member, International Biometric Society

C. Selected peer-reviewed publications (last 5 years)


Additional recent publications of importance to the field


BIOGRAPHICAL SKETCH

NAME
Xiaofei Wang, PhD

POSITION TITLE
Associate Professor of Biostatistics

eRA COMMONS USER NAME (credential, e.g., agency login)
xiaofei.wang

EDUCATION/TRAINING  (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>East China Normal University, Shanghai, China</td>
<td>BS</td>
<td>1990</td>
<td>Computer Science &amp; Psychology</td>
</tr>
<tr>
<td>Peking University, Beijing, China</td>
<td>MS</td>
<td>1993</td>
<td>Cognitive Sciences</td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill</td>
<td>MS</td>
<td>1999</td>
<td>Biostatistics</td>
</tr>
<tr>
<td>University of North Carolina at Chapel Hill</td>
<td>PhD</td>
<td>2003</td>
<td>Biostatistics</td>
</tr>
</tbody>
</table>

A. Personal Statement
I have strong research interests in developing innovative study designs and related statistical methods for the discovery and validation of biomarkers and classifiers based on high dimensional molecular, genetic and imaging data. My statistical methodology research have been funded by a NIH R03 grant on statistical methods for cancer screening, the Duke-NCSU-UNC cancer methodology P01 grant and Duke CTSA internal methodology grants. My methodology works have been published on influential statistical and medical journals. In the past years, I have also served in Duke CRTP master thesis committees and am advising master thesis projects of several Duke B&B master students as well as PhD dissertation of a UNC Biostatistics student. I am interested in teaching master and PhD level courses as well as supervising the incoming Duke B&B PhD students on their dissertations.

B. Positions, Honors and Professional Activities
Nov 2012-   Associate Professor of Biostatistics, Department of Biostatistics and Bioinformatics
Duke University Medical Center, Durham, NC
2004-Oct 2012 Assistant Professor of Biostatistics, Department of Biostatistics and Bioinformatics
Duke University Medical Center, Durham, NC
2011-present Co-Lead Statistician, Respiratory Committee,
Alliance (CALGB/NCCTG/ACOSOG)
2004-2011 Faculty Statistician, Respiratory Committee,
Cancer and Leukemia Group B (CALGB)
1996-2003 Graduate Research Assistant, Biometrics Consulting Laboratory,
Department of Biostatistics, School of Public Health,
University of North Carolina, Chapel Hill, NC
1993-1994 Lecturer, Department of Psychology
Beijing University, Beijing, China

1999 Delta Omega, Honor Society of American Public Health
2003 ENAR Student Paper Competition Award, International Biometrics Society

Advisor panel for the NCI funded research on Oncology Clinical Trial Accrual Study (OCTAS)
Treasurer for North Carolina Chapter, American Statistical Association

C. Selected Peer-reviewed Publications


BIOGRAPHICAL SKETCH

NAME
Robert F. Woolson

POSITION TITLE
Adjunct Professor of Biostatistics, Duke University; Professor Emeritus of Biostatistics, University of Iowa & MUSC

eRA COMMONS USER NAME (credential, e.g., agency login)
woolson

EDUCATION/TRAINING
(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>MM/YY</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Delaware, Newark</td>
<td>BA</td>
<td>1967</td>
<td>Mathematics (Statistics)</td>
</tr>
<tr>
<td>University of North Carolina, Chapel Hill</td>
<td>PhD</td>
<td>1972</td>
<td>Biostatistics (Epidemiology)</td>
</tr>
</tbody>
</table>

Personal Statement

I have a continuing interest in statistical methods for the design and analysis of neurological clinical trials and for epidemiologic studies. My statistical work has included the development of techniques for the analysis of longitudinal studies, including those with missing data. I have collaborated in a number of clinical research areas with longstanding interests in neurological clinical trials, especially stroke treatment clinical trials. I currently serve as consultant and adviser to the DSMB statisticians for the ALIAS and IMS-3 multi-center ischemic stroke trials. While at the University of Iowa (UI) I was Principal Investigator (PI) for the TOAST stroke trial’s Statistical and Data Management Center; was PI for several NIH statistical methodology development grants; and directed an NIMH T-32 Post-doctoral and Pre-doctoral Training Program in Psychiatric Epidemiology and Biostatistics - a program I directed for twenty years. After retirement from UI, and at MUSC I collaborated with several neurological research groups including the stroke trial group (e.g. ALIAS and IMS-3) and with the Parkinson’s disease methodology group, principally with Professor Peng Huang on methods for global statistic assessments in neurological trials. Other research work includes methodological research on informative censoring and joint modeling of survival and longitudinal data with two of my former doctoral students, Professors M. Mori (University of Oregon) and M. Jaffa (AUB).

Positions

1967 – 1968 Research Associate in Biostatistics, University of North Carolina, Chapel Hill, NC
1970 – 1972 Research Associate in Biostatistics, University of North Carolina, Chapel Hill, NC
1972 – 1977 Assistant Professor of Biostatistics, Dept. of Preventive Medicine & Environmental Health, University of Iowa, Iowa City, IA
1977 – 1981 Associate Professor of Statistics, Dept. of Statistics, University of Iowa, Iowa City, IA
1977 – 1981 Associate Professor of Biostatistics, Dept. of Preventive Medicine & Environmental Health, University of Iowa, Iowa City, IA
1981 – 1985 Professor of Biostatistics, Dept. of Preventive Medicine & Environmental Health, University of Iowa, Iowa City, IA
1981 – 2002 Professor of Statistics, Dept. of Statistics, University of Iowa, Iowa City, IA
1985 – 1999 Professor and Director of Biostatistics Division, Dept. of Preventive Medicine & Environmental Health, University of Iowa, Iowa City, IA
1999 – 2002 Professor & Head, Dept. of Biostatistics, College of Public Health, University of Iowa, Iowa City, IA
2000 – 2002 Associate Dean for Research, College of Public Health, University of Iowa, Iowa City, IA
2002 – Present Emeritus Professor, Departments of Biostatistics, Epidemiology & of Statistics, University of Iowa
2002 – 2009 Professor, Dept. of Biostatistics, Bioinformatics & Epidemiology, Medical University of South Carolina, Charleston, SC (Division of Biostatistics & Epidemiology in the Dept. of Medicine effective July 2009)

2009 – Present Emeritus Professor, Division of Biostatistics & Epidemiology, Department of Medicine, Medical University of South Carolina

2009- Present Adjunct Professor, Department of Biostatistics & Bioinformatics, Duke University, Durham, NC; Senior Statistical Consultant, Center for Health Services Research in Primary Care, Durham VA, Durham, NC

**Awards & Research Service**

1973 Delta Omega Society (Honorary Public Health Society)
1979 Foundations’ Fund for Research in Psychiatry Scholar Award—Sabbatical Award to Oxford University
1981 Mortimer Spiegelman Gold Medal Recipient, American Public Health Association
1989 (elected) Fellow, Royal Statistical Society
1990 (elected) Fellow, American Statistical Society
1999 Regents Award for Faculty Excellence, University of Iowa
2003 Centennial Distinguished Graduate Alumnus, Graduate College, UNC at Chapel Hill
1980-2012 Multiple committee appointments including: FDA Advisory Committees, NIH Study Sections

**Representative Publications (from 200 +)**

I am a PhD biostatistician with research interests in survival analysis, categorical data analysis, non-parametric analysis and statistical computing. March 2013 I joined Duke University as an assistant professor of biostatistics and bioinformatics and currently I am working on multiple clinical cancer studies. I received my PhD from the University of Iowa with thesis topic on bivariate survival analysis. Prior to the current position, from July 2011 to February 2013 I was a postdoctoral scientist of the Division of Biomedical Informatics at the University of California San Diego and working on multiple medical and biomedical research topics, from July 2010 to June 2011 I was a postdoctoral scholar of the Department of Epidemiology at the University of Iowa and serving as a biostatistician for multiple health care related projects.

Positions and Honors

Employment and Experience
2013.3 – present Assistant professor, Department of Biostatistics and Bioinformatics, Duke University
2011.7 – 2013.2 Postdoctoral scientist, Division of Biomedical Informatics, UCSD
2010.7 – 2011.6 Postdoctoral scholar, Department of Epidemiology, U. of Iowa

Awards and Honors
2012 Best Student Award Finalist, AMIA Annual Symposium
2009 Graduate Summer Fellowship, U. of Iowa
2007, 2008 Summer Merit Fellowship, AMCS Program, U. of Iowa
2003.8 – 2003.12 Graduate Fellowship, Idaho State University

Professional Societies
International Biometric Society Eastern North American Region (ENAR), American Statistical Association (ASA)

A. Publications (in reverse chronological order)
November 5, 2013

Elizabeth DeLong, PhD
7010 North Pavilion
DUMC Box 3850

Dear Liz,

I am writing in strong support of your proposal for a new PhD program in Biostatistics. I am, in particular, writing in response to one of the questions raised by the Academic Programs Committee (APC) last week.

Understanding that faculty members in your department (and throughout the School of Medicine) generally do not have a large portion of their salary fixed as compensation for teaching, they asked whether funding would be available to support the faculty time devoted teaching, mentoring and student supervision needed for a successful PhD program. I assured the APC that the School of Medicine has provided you access to ample funds for this purpose – in fact, far more than I can imagine you would need to cover faculty effort associated with the PhD program, even if the program grew substantially over time. You and I have discussed this recently, and it was documented in the confidential letter I sent last week to ask you to serve an additional five year term as Chair of Biostatistics and Bioinformatics.

In addition to this source of support, we anticipate that revenues from your successful professional Master’s degree program will also be used to support the PhD program.

I hope that this letter provides the APC, the Provost, and the Executive Committee of the Academic Council with the information requested.

Sincerely,

Nancy C. Andrews, MD, PhD
Dean, Duke University School of Medicine
Vice Chancellor for Academic Affairs
26 April, 2013

Paula McClain, Ph.D., Dean
John Klingensmith, Ph.D., Associate Dean for Academic Affairs,
The Graduate School
Duke University

Dear Paula and John,

I am writing to provide my continued strong endorsement for Liz DeLong’s proposed PhD program in Biostatistics and to assure you that I am comfortable with the revised document.

I am happy to report that the proposal for this new program has been endorsed by both our Basic Sciences and Clinical Sciences Faculty Standing Committees. It was presented to the Medical Center Executive Committee earlier this year, and unanimously supported by the voting members.

I would like to point out that Biostatistics and Bioinformatics has an outstanding faculty, well prepared for educating PhD student. In contrast to departments in other schools of the university, there is little correlation between tenured/tenure track appointments and scholarship in Biostatistics and Bioinformatics. The department is relatively young, and has very few tenure slots available. Nonetheless, it has recruited an extremely capable faculty, using long-term contracts and other mechanisms in lieu of tenured/tenure track appointments. Medical schools across the country (particularly those schools we compare ourselves to) tend to put much less emphasis on traditional tenure than other disciplines. I feel strongly that this cultural difference should not be an impediment to the formation of this important Ph.D. program.

I am very hopeful that we will be in a position to matriculate our first class in the fall of 2014. Please let me know if there is anything I can do to help this go swiftly through the remaining approval processes.

Sincerely,

Nancy C. Andrews, M.D., Ph.D.
Dean of the School of Medicine
Vice Chancellor for Academic Affairs
APPENDIX J

Excerpt from Master of Biostatistics Program Assessment Plan (SACS format)

Table 1: Program Objectives and Student Learning Outcomes

1. Objective: To develop students who acquire the analytical, biological, and communication competencies required to perform the following core functions by the time they graduate:
   Outcomes:
   A. Graduates will integrate knowledge of the scientific issues well enough to communicate with the scientific investigators, and to propose protocol changes if needed.
   B. Graduates will be able to perform standard analyses, including descriptive analyses, linear regression, analysis of variance, analysis of covariance, logistic regression, and straightforward survival analysis.
   C. Graduates will be able to identify non-standard situations requiring additional input from senior statisticians and investigators.
   D. Graduates will be able to perform/oversee data management.

2. Objective: To prepare students who desire to enter the labor force, success in finding high-quality jobs in the field.
   A. Students so identified will be invited to continue internship positions after their first semester of employment as a student.
   B. Students so identified will receive letters of recommendation for employment from internship sites.
   C. Fifty percent of students so identified will receive offers of employment from internship sites.

3. Objective: To prepare students who desire to pursue doctoral level training for admission to high-quality graduate programs.
   Outcomes:
   A. Students so identified will complete the qualifying examination without condition or re-examination at the end of year one.
   B. Students so identified will earn a grade of H in at least 75% of their coursework.
   C. Students so identified will be admitted to a PhD program within one year of completing the Master’s degree program.
**Table 2: Student Outcomes Assessment Plan**

Objective 1: To develop students who acquire the analytical, biological, and communication competencies required to perform core functions by the time they graduate.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Evidence of Outcome</th>
<th>Frequency of Collection</th>
<th>Semesters Report Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduates will integrate knowledge of the scientific issues well enough to communicate with the scientific investigators, and to propose protocol changes if needed.</td>
<td>Masters Project and Defense</td>
<td>Annually</td>
<td>Spring</td>
</tr>
<tr>
<td>Graduates will be able to perform standard analyses, including descriptive analyses, linear regression, analysis of variance, analysis of covariance, logistic regression, and straightforward survival analysis.</td>
<td>Performance in courses and qualifying examination results</td>
<td>Annually</td>
<td>Fall semester</td>
</tr>
<tr>
<td>Graduates will be able to identify non-standard situations requiring additional input from senior statisticians and investigators.</td>
<td>Employer reports</td>
<td>1 year out from graduation for each cohort</td>
<td>Spring</td>
</tr>
<tr>
<td>Graduates will be able to perform/oversee data management.</td>
<td>Employer reports</td>
<td>1 year out from graduation for each cohort</td>
<td>Spring</td>
</tr>
</tbody>
</table>
Objective 2: To prepare students who desire to enter the labor force, success in finding high-quality jobs in the field.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Evidence of Outcome</th>
<th>Frequency of Collection</th>
<th>Semesters Report Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students so identified will be invited to continue internship positions after their first semester of employment as a student.</td>
<td>Internship site reports and evaluation forms.</td>
<td>Annually</td>
<td>Fall</td>
</tr>
<tr>
<td>Students so identified will receive letters of recommendation for employment from internship sites.</td>
<td>Copies of letters in student dossiers.</td>
<td>Annually</td>
<td>Spring</td>
</tr>
<tr>
<td>Fifty percent of students so identified will receive offers of employment from internship sites.</td>
<td>Internship site reports and evaluation forms.</td>
<td>Annually</td>
<td>Summer</td>
</tr>
</tbody>
</table>

Objective 3: To prepare students who desire to pursue doctoral level training for admission to high-quality graduate programs.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Evidence of Outcome</th>
<th>Frequency of Collection</th>
<th>Semesters Report Due</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students so identified will complete the qualifying examination without condition or re-examination at the end of year one.</td>
<td>Qualifying examination report</td>
<td>Annually</td>
<td>Fall</td>
</tr>
<tr>
<td>Students so identified will earn a grade of H in at least 75% of their coursework.</td>
<td>Student academic record and dossiers.</td>
<td>Annually</td>
<td>Fall</td>
</tr>
<tr>
<td>Students so identified will be admitted to a PhD program within one year of completing the Master’s degree program.</td>
<td>Correspondence from admitting program.</td>
<td>Annually</td>
<td>Fall</td>
</tr>
</tbody>
</table>
Table 3: Partial Analytical Report - AY2013

(At this time, we are six months in advance of the graduation of our first cohort. Hence the analytical report is based primarily on the end-of-semester debriefings attended by faculty members and program staff.)

<table>
<thead>
<tr>
<th>Selected Outcomes</th>
<th>Evidence collected</th>
<th>Finding and evaluation</th>
<th>Resultant action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to perform standard analyses, including descriptive analyses, linear regression, analysis of variance, analysis of covariance, logistic regression, and straightforward survival analysis.</td>
<td>Performance in courses and internship reports.</td>
<td>Faculty reports of performance during the internship experience have been positive to date. The only source of concern was from one of the internships in the private sector, noting that our training in computer programming could be upgraded.</td>
<td>1. First year computer programming course has been restructured.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2. Students interested in obtaining SAS certification may apply for funding from department for training fee.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. The restructured programming course will be revisited at the end of the academic year for consideration of expansion to 2 credits and/or implementation of competency test.</td>
</tr>
<tr>
<td>Complete the qualifying examination without condition or re-examination at the end of year one.</td>
<td>Qualifying exam reports.</td>
<td>All of our full-time students passed the qualifying exam in August, 2012.</td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX K

### Grants on which B&B Faculty are Principal Investigators

<table>
<thead>
<tr>
<th>Name</th>
<th>Project Role</th>
<th>Sponsor</th>
<th>Prime Sponsor</th>
<th>Annual Directs</th>
<th>Project Start</th>
<th>Project End</th>
<th>PI Name</th>
<th>Proposal Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anstrom, Kevin J.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>375,223</td>
<td>6/8/12</td>
<td>4/30/17</td>
<td>Anstrom, Kevin J.</td>
<td>GUIDing Evidence Based Therapy Using Biomarker Intensified Treatment (DCC)</td>
</tr>
<tr>
<td>Anstrom, Kevin J.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>3,826,431</td>
<td>1/1/12</td>
<td>12/31/18</td>
<td>Anstrom, Kevin J.</td>
<td>Heart Failure Clinical Research Network Coordinating Center</td>
</tr>
<tr>
<td>Anstrom, Kevin J.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>4,660,745</td>
<td>5/1/05</td>
<td>4/30/14</td>
<td>Anstrom, Kevin J.</td>
<td>Idiopathic Pulmonary Fibrosis Clinical Research Network</td>
</tr>
<tr>
<td>Anstrom, Kevin J.</td>
<td>PI</td>
<td>UCSF</td>
<td>NIH</td>
<td>189,587</td>
<td>8/1/13</td>
<td>5/31/16</td>
<td>Anstrom, Kevin J.</td>
<td>IPF Wrap</td>
</tr>
<tr>
<td>Barnhart, Huiman x.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>907,495</td>
<td>9/3/13</td>
<td>6/30/18</td>
<td>Barnhart, Huiman x.</td>
<td>Coordinating Center for the Drug Induced Liver Injury Network (DILIN)</td>
</tr>
<tr>
<td>Chan, Cliburn C.</td>
<td>Core Director</td>
<td>NIH</td>
<td>NIH</td>
<td>100,221</td>
<td>7/15/10</td>
<td>6/30/15</td>
<td>Chan, Cliburn C.</td>
<td>Biostatistics and Computational Biology Core</td>
</tr>
<tr>
<td>Chan, Cliburn C.</td>
<td>PI</td>
<td>Leiden University</td>
<td>Wallace Couler</td>
<td>167,349</td>
<td>4/1/12</td>
<td>3/31/15</td>
<td>Chan, Cliburn C.</td>
<td>Standardization of Immune Monitoring - phase II</td>
</tr>
<tr>
<td>Delong, Elizabeth</td>
<td>Consortium PI</td>
<td>Tufts University</td>
<td>PCORI</td>
<td>133,667</td>
<td>6/1/13</td>
<td>5/31/16</td>
<td>Delong, Elizabeth</td>
<td>Integrating Causal Inference, Evidence Synthesis, and Research Prioritization Methods</td>
</tr>
<tr>
<td>Erkanli, Alaattin</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>714,549</td>
<td>4/1/11</td>
<td>3/31/16</td>
<td>Ramanujam, Nirmala</td>
<td>A Novel Optical Spectral Imaging System for Rapid Imaging of Breast Tumor Margins</td>
</tr>
<tr>
<td>George, Stephen L.</td>
<td>PI</td>
<td>UNC-CH</td>
<td>NIH</td>
<td>-</td>
<td>3/1/15</td>
<td>3/31/15</td>
<td>George, Stephen L.</td>
<td>Statistical Methods for Cancer Clinical Trials</td>
</tr>
<tr>
<td>Gordan, Raluca M.</td>
<td>PI</td>
<td>PhRMA Foundation</td>
<td>PhRMA Foundat</td>
<td>60,000</td>
<td>1/1/13</td>
<td>12/31/13</td>
<td>Gordan, Raluca M.</td>
<td>Understanding the complex interplay among members of a transcription factor family</td>
</tr>
<tr>
<td>Gordan, Raluca M.</td>
<td>PI</td>
<td>March of Dimes</td>
<td>March of Dimes</td>
<td>136,364</td>
<td>2/1/13</td>
<td>1/31/15</td>
<td>Gordan, Raluca M.</td>
<td>Using high-throughput protein binding microarray data to identify DNA sequence polymorphisms</td>
</tr>
<tr>
<td>Halabi, Susan</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>148,920</td>
<td>7/6/11</td>
<td>5/31/15</td>
<td>Halabi, Susan</td>
<td>Prognostic Models of Clinical Outcomes In Men With Castration Resistant Prostate Cancer</td>
</tr>
<tr>
<td>Halabi, Susan</td>
<td>PI</td>
<td>Cedars Sinai Medical Center</td>
<td>NIH</td>
<td>26,367</td>
<td>9/12/12</td>
<td>6/30/14</td>
<td>Halabi, Susan</td>
<td>Expression Profiling of Renal Cell Carcinoma Utilizing Tissue from CALGB 90206</td>
</tr>
<tr>
<td>Halabi, Susan</td>
<td>PI</td>
<td>Mayo Clinic</td>
<td>Mayo Clinic</td>
<td>10,000</td>
<td>6/1/13</td>
<td>5/31/14</td>
<td>Halabi, Susan</td>
<td>Statistical Analysis of Overall Survival in Complex Situations of Phase II Clinical Trials</td>
</tr>
<tr>
<td>Jung, Sin-Ho</td>
<td>PI</td>
<td>Roswell Park Cancer Institute</td>
<td>NIH</td>
<td>23,147</td>
<td>2/1/12</td>
<td>1/31/14</td>
<td>Jung, Sin-Ho</td>
<td>Diet Change Among Prostate Cancer Patients Under Expectant Management</td>
</tr>
<tr>
<td>Jung, Sin-Ho</td>
<td>PI</td>
<td>UNC-CH</td>
<td>NIH</td>
<td>18,901</td>
<td>4/1/10</td>
<td>3/31/15</td>
<td>Jung, Sin-Ho</td>
<td>Methods for Missing and Auxiliary Data in Cancer Clinical Trials</td>
</tr>
<tr>
<td>Lee, Kerry L.</td>
<td>PI</td>
<td>UNC-CH</td>
<td>NIH</td>
<td>39,406</td>
<td>4/1/10</td>
<td>3/31/15</td>
<td>Lee, Kerry L.</td>
<td>Methods for Pharmacogenomics and Individualized Therapy Trials</td>
</tr>
<tr>
<td>Lee, Kerry L.</td>
<td>PI</td>
<td>Eli Lilly and Company</td>
<td>Eli Lilly and Company</td>
<td>324,132</td>
<td>12/1/08</td>
<td>4/30/14</td>
<td>Lee, Kerry L.</td>
<td>CABANA-AF Trial</td>
</tr>
<tr>
<td>Lee, Kerry L.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>560,496</td>
<td>9/29/09</td>
<td>11/30/14</td>
<td>Lee, Kerry L.</td>
<td>PROMISE Trial: SDCC</td>
</tr>
<tr>
<td>Lee, Kerry L.</td>
<td>PI</td>
<td>Mount Sinai School of Medicine</td>
<td>NIH</td>
<td>172,178</td>
<td>3/1/13</td>
<td>2/28/14</td>
<td>Lee, Kerry L.</td>
<td>TACT-Trial to Assess Chelation Therapy Data Coordinating Center NCE</td>
</tr>
<tr>
<td>Li, Zhiguo</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>169,670</td>
<td>9/30/09</td>
<td>6/30/14</td>
<td>Li, Zhiguo</td>
<td>Biostatistical Core</td>
</tr>
<tr>
<td>Li, Zhiguo</td>
<td>PI</td>
<td>UNC-CH</td>
<td>NIH</td>
<td>9,845</td>
<td>4/1/10</td>
<td>3/31/15</td>
<td>Li, Zhiguo</td>
<td>Methods for Discovery and Analysis of Dynamic Treatment Regimes</td>
</tr>
<tr>
<td>Martin, Barbara E.</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>158,599</td>
<td>9/6/11</td>
<td>6/30/15</td>
<td>Martin, Barbara E.</td>
<td>Statistical Models to Investigate Long-Distance QTL Transcription Regulation</td>
</tr>
<tr>
<td>Niedzwiecki, Donna</td>
<td>PI</td>
<td>Dana-Farber Cancer Institute</td>
<td>NIH</td>
<td>18,296</td>
<td>7/1/10</td>
<td>5/31/14</td>
<td>Niedzwiecki, Donna</td>
<td>The Influence of Diet and Lifestyle on Patients with Advanced Colorectal Cancer</td>
</tr>
<tr>
<td>O'Brien, Sean M.</td>
<td>PI</td>
<td>STS</td>
<td>STS</td>
<td>40,118</td>
<td>12/31/13</td>
<td>12/31/13</td>
<td>O'Brien, Sean M.</td>
<td>Linking the Congenital Heart Surgery Database of the STS with the CHSS Database</td>
</tr>
<tr>
<td>O'Brien, Sean M.</td>
<td>Consortium PI</td>
<td>North Carolina State University</td>
<td>NIH</td>
<td>97,648</td>
<td>6/1/13</td>
<td>5/31/14</td>
<td>O'Brien, Sean M.</td>
<td>Statistical Methods for Complex Data in Cardiovascular Disease</td>
</tr>
<tr>
<td>O'Brien, Sean M.</td>
<td>PD/PI</td>
<td>NIH</td>
<td>NIH</td>
<td>2,677,897</td>
<td>7/22/11</td>
<td>10/31/18</td>
<td>O'Brien, Sean M.</td>
<td>The ISCHEMIA Trial - SDCC</td>
</tr>
<tr>
<td>Owarz, Kouroso</td>
<td>PI</td>
<td>NIH</td>
<td>NIH</td>
<td>70,066</td>
<td>8/1/12</td>
<td>7/31/15</td>
<td>Owarz, Kouroso</td>
<td>Genome-Wide Predictors of Treatment-related Toxicities</td>
</tr>
<tr>
<td>Owarz, Kouroso</td>
<td>PI</td>
<td>UNC-CH</td>
<td>NIH</td>
<td>16,259</td>
<td>4/1/10</td>
<td>5/31/15</td>
<td>Owarz, Kouroso</td>
<td>Computational Resources and Dissemination Core</td>
</tr>
<tr>
<td>Owarz, Kouroso</td>
<td>PI</td>
<td>Roswell Park Cancer Institute</td>
<td>NIH</td>
<td>25,732</td>
<td>1/1/13</td>
<td>12/31/13</td>
<td>Owarz, Kouroso</td>
<td>Bioinformatics</td>
</tr>
<tr>
<td>Owarz, Kouroso</td>
<td>PI</td>
<td>University of Chicago</td>
<td>NIH</td>
<td>52,771</td>
<td>7/16/10</td>
<td>6/30/15</td>
<td>Owarz, Kouroso</td>
<td>PAAR Pharmacogenomics of Anticancer Agents Research Group</td>
</tr>
<tr>
<td>Pang, Herbert H.</td>
<td>PD/PI</td>
<td>NIH</td>
<td>NIH</td>
<td>131,937</td>
<td>8/1/13</td>
<td>5/31/15</td>
<td>Pang, Herbert H.</td>
<td>Translational meta-analysis for elderly lung cancer patients</td>
</tr>
<tr>
<td>Pieper, Carl F.</td>
<td>Core Leader</td>
<td>NIH</td>
<td>NIH</td>
<td>-</td>
<td>9/15/06</td>
<td>6/30/16</td>
<td>Pieper, Carl F.</td>
<td>RCI</td>
</tr>
<tr>
<td>Reddy, Timothy E.</td>
<td>PI</td>
<td>Northwestern University</td>
<td>NIH</td>
<td>151,785</td>
<td>7/11/13</td>
<td>5/31/14</td>
<td>Reddy, Timothy E.</td>
<td>Genetics and Evolution of Fetal Human Fat Accretion During Development</td>
</tr>
<tr>
<td>Roy Choudhury, Kingsh Co-PI</td>
<td>PI</td>
<td>North Carolina Biotechnology Center</td>
<td>NIH</td>
<td>300,000</td>
<td>2/1/13</td>
<td>1/31/15</td>
<td>Roy Choudhury, Kingsh Co-PI</td>
<td>Imaging, Modeling and Modulating Stomal Cell Roles in Tumor Growth and Therapy</td>
</tr>
<tr>
<td>Wang, Xiaofei F.</td>
<td>PD/PI</td>
<td>NIH</td>
<td>NIH</td>
<td>131,937</td>
<td>8/1/13</td>
<td>5/31/15</td>
<td>Wang, Xiaofei F.</td>
<td>Translational meta-analysis for elderly lung cancer patients</td>
</tr>
</tbody>
</table>

**APPENDIX K**

**Grants on which B&B Faculty are Principal Investigators**