Biostatistics Departmental Retreat
Burkshire Conference Center
Saturday, February 27, 1999
8:30 AM - 3:00 PM

MINUTES

Present:	**Faculty**:	Scott Zeger, Helen Abbey, Karen Bandeen-Roche, Ron Brookmeyer, Frank Curriero, Marie Diener-West, Mitch Gail, Steve Goodman, Rafael Irizarry, Joanne Katz, Tony Lachenbruch, Kung-Yee Liang, Lucy Mead, Bill Rising, Charles Rohde, Daniel Scharfstein, Matthew Tayback, Rick Thompson, Jim Tonascia, Mei-Cheng Wang, Colin Wu
	Students:	Natalie Blades, Liz Garrett, Wei-Ting Hwang
	Staff:	Mary Joy Argo, Chris McCullough, Debra Moffitt, Jiong Yang

1) State of the Department Report

Dr. Zeger opened the retreat by welcoming the participants and thanking them for taking a Saturday to discuss strategic issues facing the Department of Biostatistics. Dr. Zeger began by summarizing accomplishments during the past year. These include:

- The hiring in 1998-99 of new faculty and staff: Rafael Irizarry and Colin McCulloch as assistant professors; Richard Thompson as research associate; Bill Rising and Frank Curriero as assistant scientists; Jiong Yang as the Department's computer programmer; the participation of Tony Lachenbruch and Larry Magder as adjunct faculty teaching Montgomery County courses;

- A successful recruitment year in which Constantine Frangakis from Harvard's Department of Statistics, Karl Broman from the University of California-Berkeley, and possibly others, will be joining us in 1999-2000;

- A successful research year as evidenced by 40 new papers by departmental faculty and a new book contract signed by Alan Ross;

- The implementation of the new introductory course sequences: 611-12 by Ron Brookmeyer; 621-24 by Marie Diener-West, Jim Tonascia, and Scott Zeger; the new social science course 140.658 taught jointly by Karen Bandeen-Roche and Bill Eaton of Mental Hygiene; and an advanced computing course 140.778 taught by Colin McCulloch for our PhD students;

- The creation of the Biostatistics Center, which in its eighteen-month existence has already served some 50 investigators from within Johns Hopkins and several external clients, including the states of Minnesota, Maryland, and Wisconsin in their tobacco suits, Covance, and Meridian

Technologies;

- The assignment of new space to the Department, which will provide approximately twelve new offices for faculty, staff, and members of the Biostatistics Center;

- A gift of $500,000 over ten years from alumnus Frank Hurley (PhD '71) of Quintiles;

- The creation of a departmental homepage with extensive information about the faculty, students, staff, history, courses, and other aspects of our daily operations. The homepage is used regularly by prospective students and faculty candidates and has made the Department considerably more competitive.

In addition to this partial list of accomplishments, Dr. Zeger raised several issues for further consideration during the retreat, including:

- a reduction in the number and quality of applicants to the PhD program, leading to a few smaller PhD entering classes;
- the difficulty some PhD students are having with the Department's advanced classes;
- the difficulty some students are experiencing in finishing their PhD research in 4.5-5 years;
- declining enrollment in the Department's introductory sequences, partly due to the declining numbers of PhD and MPH students;
- the growing dependence of the Department on grants in which our faculty are principal investigators and the possibility that biostatistics grants will once again become difficult to obtain with upcoming changes at NIH.

Chris McCullough and Mary Joy Argo led the attendees through a series of tables summarizing the financial and academic status of the Department to date. The main points of the discussion were:

- Departmental general funds are projected to be slightly higher in fiscal year 2000 than in fiscal year 1999. But this is the result of a one-time addition from non-tuition sources of approximately $1 million to the TAM to offset major increases in central expenditures in support of academic programs.

- A decrease in Biostatistics enrollments from 2,340 to 2,018 over the last academic year, and a decrease in credits earned from 9,495 to 8,220.

- Sponsored projects to the Department have grown from $650,000 to $2 million since 1996. This growth represents new methodologic grants to teams led by Mei-Cheng Wang, Kung-Yee Liang, Ron Brookmeyer, and Scott Zeger. The high degree of dependence on these sources of funds could place the Department at risk starting around the year 2001, if NIH once again makes biostatistical grants difficult to obtain.

- A reduction in tuition allocation for courses taught in the full-time program, from $928,000 to $836,000 over the last year. This is largely the result of a reduced number of PhD, MPH, and MHS students in our courses. It is projected that some additional drop is likely because of

decreasing enrollment in the School, as well as the revision to our introductory sequences that may have the effect of decreased credits being earned in the Department of Biostatistics.

- A growth in the full-time professorial faculty from six professors, four associate professors, and three assistant professors, to seven professors, two associate professors, and between five and seven assistant professors (depending upon the number of offers accepted).

- A growth in the number of courses offered by the Department from 41 in 1997-98 to 47 in the current academic year. This growth includes the following new courses: 610, 611-12, 621-624, 632 (SAS), 658 (social science statistics), and 778 (advanced computing). It also includes the discontinuation of courses 601-604.

- A delay in the receipt of applications, which are now peaking in February-March rather than January-February.

2) Statistical Education for Public Health Professionals and Scientists

Marie Diener-West led this discussion, which began with a review of our new course sequences. Ron Brookmeyer first reported on the new two-term, six-credit sequence 611-612 (Statistical Reasoning in Public Health). His objective in designing this sequence was for the curriculum to be just as rigorous as 621-24 but with a focus on the main statistical ideas and on preparing students to read the biomedical literature, rather than conducting or participating in data analysis. The sequence covers most of the major statistical ideas and methods used in the literature, from basic methods of inference through logistic and survival regression analyses. Special attention was given to make the homeworks and exams challenging enough so that students would not view this sequence as an "easy way out" of their statistics requirement. In this regard, Ron felt that special students make it more difficult to achieve the desired level of rigorous thinking required of our School's degree candidates.

Marie Diener-West summarized the new 621-24 sequence, which she also felt was going well. With the availability of the new 611-12 sequence, students in 621-24 are more homogeneous; all are attempting to develop data analytic skills. Statistical computing using STATA has been integrated into the course. In this regard, Marie reported some initial difficulty, before students gained familiarity with the software. It was suggested that we provide STATA scripts to students early in the course, so that they could focus on the statistical reasoning and methods and not on their implementation in software.

Ron pointed out that several students in 611-12 were interested in continuing on with additional statistics courses. There was some discussion of an alternate approach to introductory biostatistics whereby students start with 611-12, to cover the main statistical ideas, and then move on to developing data analysis skills in two additional terms. One possibility was to restructure the statistical computing courses to have more statistical content; students would thus gain more advanced data analysis skills in follow-up to 611-12. Another idea was to have all students take 611-12 and then design two new terms that would cover the additional materials in 621-24, which is not currently covered in 611-12. This would be analogous to what Epidemiology does with Epi 1,

followed by Epi 2, 3, and 4. This strategy would be one way to better coordinate with the Department of Epidemiology, but it is not realistic for the short term given our very substantial investment in the new sequences nor is it clear that a two-term course to follow 611-12 would be as effective as the four terms we now have with 621-24.

It was decided that the Department would convene a focus group of students in the 611-12 and 621-24 sequences to obtain their feedback. Mary Joy Argo will organize these meetings over the next few months. It was also decided that we meet with the academic coordinators and/or chairs of the various departments to learn more about their thoughts on our new courses.

Finally, Karen Bandeen-Roche reported on her new social science statistics course – a two-term sequence taught jointly with Bill Eaton of Mental Hygiene. There were roughly 40 students in the first term and 20 in the second. Students who took the course reported that the interchange between Karen and Bill was highly valued. They indicated that supplementary discussions specific to the Biostatistics and Mental Hygiene students were helpful when provided and could be usefully expanded in the future. Overall, this model of jointly teaching second-year doctoral courses was thought to have substantial merit.

The discussion then turned to a previous proposal to replicate this idea in the area of epidemiology. We would introduce a new two-term sequence for second year doctoral students on statistical methods for epidemiology. Steve Goodman described a planning process about to begin in the Department of Epidemiology to consider a second-year methods sequence, part of which would include statistical methods. He was supportive of the idea of having a statistics component of this second-year methodology sequence, to be taught jointly by faculty from Biostatistics and Epidemiology, following the social sciences model. In this way, we could avoid having advanced statistical methods courses for epidemiologists taught by both the Departments of Biostatistics and Epidemiology. Further discussion of this idea can take place in the committee meetings to be chaired by Steve; Ron Brookmeyer will be participating.

3) Future of Biostatistics

The discussion then turned to where the field of biostatistics is going in the future, and what it will take for our own department to achieve its maximum potential within the Hopkins context. The major trends in biomedical and public health research include the following:

- The genomics revolution is producing new scientific insights and vast amounts of DNA sequence and marker data that will require new study designs and methods of analysis.

- New technologies for visualization and process monitoring are leading to larger and more complex data sets. Examples include neuroimages, visual fields, geographic data on morbidity and mortality, and multivariate longitudinal data.

- Large cohort studies such as the Women's Health Initiative will produce large, complex longitudinal data sets and numerous follow-up studies that will combine original data with specialized information collected to address particular questions. New designs and analytic

techniques will be necessary.

- An explosion in information technologies, making possible more complex modeling and estimation procedures, such as Markov Chain Monte Carlo for Bayesian models.

- Distributed computing in which large numbers of individuals work on a common project across the internet. The LYNEX operating system development is an example where more than 100 programmers around the world donate their time to the development of a highly-competitive (and free) alternative operating system to Windows. A second example is the epidemiology curriculum being developed over the internet, whereby a large number of faculty at many universities each contribute one lecture or part of a lecture.

- Emphasis on surrogate endpoints and biomarkers in clinical research, resulting from increased understanding of the mechanisms underlying the efficacy and safety of new drugs.

- Increased inequality in the distribution of the world's wealth, leading to two classes of public health problems in the developing and industrialized worlds respectively.

4) PhD Program:

Kung-Yee Liang led a discussion of strengths and weaknesses of our current PhD program. There was general agreement that the success of our Department is intimately tied to the quality of our PhD students and graduates. Natalie Blades, Wei-Ting Hwang, and Liz Garrett discussed their impressions of the current program and ways in which the program could be made more attractive to applicants. Liz Garrett expressed gratitude that several student concerns raised at last year's retreat were successfully addressed during the past academic year. She also indicated that the opportunity to apprentice with faculty on research projects was most important to her, given her interest in working at the interface of statistics and biomedical science. There was some discussion of the value of offering real analysis and advanced probability in the first year. Most felt that this was a good idea. There was a consensus that we should consider joining with the Department of Mathematical Sciences in teaching introductory probability and statistics so that our students and theirs could interact more and so that the small introductory courses in each of our two departments could be combined to save faculty time. Scott Zeger indicated that he has spoken with faculty in Mathematical Sciences and will continue to explore possibilities for cooperation.

The discussion turned to how to attract more excellent students like Patrick Heagerty and Paul Rathouz, two of our best graduates in recent years. It was noted that both Patrick and Paul had experience after their undergraduate training prior to returning to graduate school. Both had been Peace Corps volunteers. Patrick had taught mathematics and earned a master's degree at SUNY-Albany, while Paul had earned a master's degree from the University of North Carolina. It was agreed that maturity increases a student's chance for success. Possible sources of such students include returned Peace Corps volunteers, MD/PhD students, master's graduates from other programs in other quantitative fields, persons working in industry looking to return to graduate training (ie, Mark Schactman). Additional discussions about how to optimally pursue the 2000-2001 matriculating class will be continued over the next few months.

5) Measuring Our Success

The final discussion addressed the question: how will we measure our success over the next 5-10 years? The first and perhaps most important criterion is that we regularly publish important statistical methodologic and substantive research in the best journals. There was agreement that by finding solutions to important statistical and substantive problems, we will improve the fields of biostatistics and public health and will attract the best new faculty and students over the long term.

Another measure of success is our continued ability to attract statistical methods grants to the Department. Doing so will require a high degree of productivity on current grants so that they can be renewed and others initiated. A related measure of success is the use by other statisticians and substantive researchers of statistical ideas and methods developed by our faculty. Creating and disseminating software for our new methods is a key component here. Our success can also be reflected in the quality of our PhD students and graduates. The influence of our research can be magnified when students and graduates take on and solve related problems.

SLZ/mja