Report of the October 28, 2000 Retreat
Johns Hopkins Department of Biostatistics

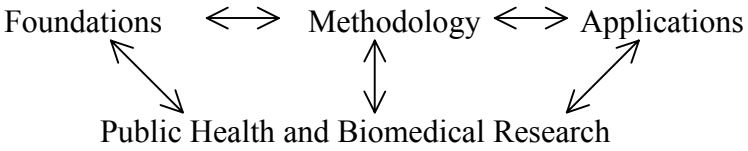
This report summarizes themes discussed at the October 28, 2000 retreat of the Johns Hopkins Department of Biostatistics at Evergreen House and in faculty meetings over the last several months. A list of retreat attendees is in Table 1. The major ideas have been organized into separate sections. By way of background, the report begins with an overview of the Johns Hopkins perspective (as understood by the current chair) on the field of biostatistics.

Hopkins Perspective on Biostatistics:

Biostatistics comprises the reasoning and methods for using data as evidence to address public health and biomedical questions. It is an approach and a set of tools for designing studies and for quantifying the resulting evidence, for quantifying what we believe, and for making decisions.

At Johns Hopkins Department of Biostatistics, research is characterized by a commitment to statistical science, its foundations and methods, as well as the application of statistical science to the solution of public health and biomedical problems. As indicated in the two-way arrows in Figure 1, research on foundations informs and is informed by methods research, which in turns benefits and is benefited by statistical applications. To be excellent, biostatistical research must be built on a foundation of first-rate public health and biomedical research, like that which occurs at Johns Hopkins.

Figure 1: Biostatistics Research at the Johns Hopkins University Department of Biostatistics



Research on foundations has as its goal the development of better strategies, or ways of reasoning, for empirical research. For example, past chair William Cochran demonstrated how observational studies can be used to draw inferences about the causal effect of a treatment on a health outcome. Jerry Cornfield showed how case control studies can be used to draw valid inferences about parameters in prospective models. Today, Richard Royall is leading a transition in statistical reasoning from decision methods (*p*-values, tests of hypotheses) toward likelihood methods, which quantify scientific evidence.

Research on statistical methodology has as its goal the creation of new strategies for drawing inferences from data. To illustrate, Ron Brookmeyer and Mitch Gail developed the methodology used to monitor and project the size of the US AIDS epidemic; Kung-Yee Liang, Mei-Cheng Wang, and Scott Zeger developed methods for regression analysis with correlated responses.

Dan Scharfstein and colleagues have developed graphical techniques for assessing the possible impact of missing data in clinical trials and observational studies. Kung-Yee Liang, Karl Broman, and Giovanni Parmigiani are developing new techniques to find disease genes.

Biostatistics also includes research on important substantive questions. For example, Francesca Dominici and colleagues have used multiple national databases to determine the effects of air pollution on mortality across the 90 largest American cities. Marie Diener-West, Jim Tonascia, Steve Piantadosi, and others have led or collaborated in clinical trials of new therapeutic treatments.

Throughout its history and today, the definition of biostatistics has expanded to include foundations, methodology, and applications. The faculty's commitment to this inclusive perspective and the support of the School's administration and faculty are two of the intangible yet critical components of the Department's current and future success.

Biostatistical research involves collaborations with health researchers. A lifeblood for the Department is continued, strong relationships with leading substantive experts at Johns Hopkins. In specialties like epidemiology and oncology, there are other groups of primarily-appointed statisticians. The University's Department of Biostatistics will be diminished if it loses contact with these key areas of biomedical research. The long-term success of the Department, therefore, depends upon building and maintaining mutually-supportive relationships between the groups so that primary and joint faculty members have continual exposure to leading problems in as diverse a set of specialties as possible.

Biostatistics at Johns Hopkins Health Institutions:

For half a century, the Department of Biostatistics was the only concentration of statistical expertise at the East Baltimore campus. (Until the mid-1970s, the University also had a department of statistics, which collaborated closely with Biostatistics.) Over the last few decades, there has been an enormous increase in the demand for statisticians as a result of growth in NIH-sponsored research. The Department of Biostatistics decided to keep a balance of research with statistical education and therefore, did not grow in size to meet all the needs of NIH research grants. As a result, clusters of biostatisticians have been hired in multiple departments across the Health Institutions, including Epidemiology (clinical trials, AIDS) and Oncology. Each of these groups now has a critical mass of statisticians. Small groups of statistical expertise also exist in the Departments of International Health and Population and Family Health Sciences. Similar groups are being considered by the Department of Urology and by the Kennedy Krieger Institute. In addition, the School of Medicine has created the Institute for Human Genetics, chaired by Aravinda Chakravarti, who plans to add several statistical geneticists and genetic epidemiologists.

The co-trends of increasing numbers of biostatisticians and a decentralized organization is likely to continue if NIH funding continues to grow, producing demands for statistical collaborators and consultants. These trends pose two questions, which the Department has addressed over the last few months:

- How can biostatistical faculty at Johns Hopkins best organize themselves to promote foundations, methods, and applications research to advance biostatistics and ultimately, the public's health?

- What roles should the Department of Biostatistics play over the next decade?

Role of the Department of Biostatistics:

Discussions over the last year, culminating in the retreat, have identified the following priority missions for the Department of Biostatistics:

- conduct of original research on important biostatistical problems across the spectrum: foundations ↔methodology ↔ applications;
- responsibility for Johns Hopkins University's PhD and master's programs in biostatistics;
- leadership of biostatistical education for public health/biomedical scientists and professionals at Johns Hopkins;
- participation in other current and future educational programs involving substantial statistical reasoning, such as quantitative genetics, bioinformatics, and clinical investigations;
- facilitation of biomedical and public health research that depends on statistical collaboration or consultation.

Organizing Biostatisticians to Promote the Discipline and the Public's Health:

Given continued growth and decentralization of biostatistical research and applications across the University, the faculty have also discussed how best to organize themselves to promote the discipline and to make it maximally useful in advancing health.

In our current arrangement, the multiple statistics groups develop as each chooses to meet its local opportunities and obligations and to fulfill individual faculty aspirations. Collaborations among groups occur regularly, as they are mutually beneficial. So long as the groups avoid major competitions for resources, particularly for grants and new faculty members, this arrangement has and can continue to succeed.

It has been suggested that we consider an organization like the Oncology Center. The Department of Oncology is responsible for training. The Oncology Center, built with the Department of Oncology at its core, expands the Oncology research program by including faculty across many departments. It clearly makes Hopkins Oncology and the affiliated faculty more influential in the field. A Biostatistics Center could jointly organize working groups, the seminar series, computing infrastructure, and possibly consulting and clinical trials coordinating center services. It was noted that informal cooperation tends to work better than formal organizations, particularly absent core funding such as that which exists for the Oncology Center. The success of the less formal "Field of Statistics" at Cornell was offered as an example.

Working Groups:

It was proposed that Hopkins Biostatistics research productivity and visibility can be enhanced through loose affiliations of faculty with common research goals, which we currently call "working groups." The main objective of a group would be to stimulate, collaborate on, and disseminate results of research in a particular subspecialty.

It is recognized that research actually takes place individually or in smaller groups. The goal of a working group would be to facilitate multiple individual or collaborative research efforts in a particular area and to more effectively disseminate their results. An effective working group might:

- organize an informal seminar/discussion group;
- fund a methodologic grant to support several investigators with related interests;
- create common software to implement research methods;
- disseminate information about the group through:
    - research articles
    - an active web page;
    - an annual public workshop/meeting;
    - books.

Areas where we might continue existing working groups or create new ones include:

- statistical methods for longitudinal studies;
- statistical genetics;
- foundations of inference;
- environmental statistics;
- Bayesian biostatistics;
- biostatistical practice and education.

One open question is how the Department might facilitate faculty who choose to work through a group.

Information Technologies:

We have made substantial progress over the last few years in our use of information technologies. Student applications to the Department can now be filed electronically. The Department has a web presence that provides substantial information about our work to prospective students and faculty. Many of our courses have extensive web-based educational materials. Biostatistics 610 and 611 (and soon 612) are taught entirely via the web. We have a high-speed network (100 megabits per second) to which all primarily-appointed faculty and students are connected. A large SUN server is on order, along with ten additional terminals to increase our computing power and provide greater access to the shared resource. We have a full-time computer scientist to manage our computing environment. All of these gains have been achieved quite recently.

Despite the gains, however, the most serious concerns raised by retreat participants were about the Department's information technologies. There was consensus that although we are an

information-centered discipline, our use of modern information technologies for departmental administration, research, and teaching needs to be improved. Steps for the near future might include:

- clarifying our goals with respect to: administrative computing; statistical computing research; web-based technologies for research and teaching;
- "working" on the web, wherever possible, so that we upgrade our collective understanding of its potential for research, teaching, and administration;
- hiring a statistical computing expert (like Anthony Rossini at the University of Washington or Mike Meyer at Carnegie Mellon) who can help improve the statistical computing environment;
- involving Computer Science undergraduate/graduate students from Homewood in our departmental research programs;
- collaborating with computer scientists in the School of Medicine to create a common, higher-quality computing environment.

The Biostatistics Information Technology (BIT) Committee is now responsible for prioritizing the steps we should take. The Department Chair is committed to making this our last retreat where IT limitations are identified as a departmental weakness.

Strengths and Weaknesses:

The retreat participants created a list of the Department's strengths and weaknesses, as summarized in Table 2. The text below provides an overview of some of the items discussed at greater length.

What's Working: Among the strengths singled out was the quality and diversity of the Department's faculty, staff, and students and the collegial and intellectual environment they have collectively created. What is particularly important about the environment is the commitment of faculty to statistical science as an intellectual force and the resulting definition of biostatistics to include: foundations $\leftrightarrow$ methods $\leftrightarrow$ applications. Hand in hand with the collegial environment is a commitment by each Department member to achieve excellence. We have never and cannot in the future lower our expectations for what we can contribute to statistical science and public health.

The Johns Hopkins Health Institutions environment was also singled out as key to our success. Faculty and students are surrounded by research opportunities and the highest-quality colleagues in health specialties with whom to collaborate.

The commitment of the School of Public Health -- and in particular, its Dean, Al Sommer -- to the centrality of biostatistics was also singled out.

The Department's tradition of hiring the best available young faculty and protecting them for a period of time so they can establish research programs is a tradition that strengthens us. Both faculty and students indicated there is supportive competition pushing each of us to achieve our potential.

In addition to the research environment, the faculty identified the teaching environment as key. Faculty are enthusiastic our doctoral and master's programs in biostatistics, but equally about our teaching of biostatistics to health scientists and professionals. A unique strength of the Hopkins Department of Biostatistics is the opportunity to teach non-departmental students who go on to leadership health positions around the world.

Where Work Is Needed: The most critical short-term problem faced by the Department is information systems. We need to incorporate modern, web-based technologies into the everyday workings of the Department. We need reliable and accessible systems that are competitive with those available to departments of statistics and biostatistics. We likely need to build collaborations with computer science students and faculty in the University to strengthen ourselves.

Given our current size, the Department cannot excel in all areas of biostatistics. For example, it is difficult for us to respond to opportunities in bioinformatics/genomics without having that subspecialty dominate the Department. Addressing this limitation will require innovative collaborations with statisticians, computer scientists, and other quantitatively-oriented faculty and students across the University.

Some colleagues believe that the teaching of biostatistics need not necessarily be the responsibility of the Department of Biostatistics. Concern was expressed that introductory courses are large and perceived by some to be less effective because of their size. The department must continue to offer excellent courses, whether small or large.

Research Opportunities:

Table 3 summarizes the research opportunities discussed in greatest detail at the retreat. Several of the topics are currently well-represented in the Department. These include design and analysis of data from cohort studies, statistical genetics, foundations of inference, and clinical trials methodology. In the area of statistical genetics, there was consensus that we should continue to work both in population and molecular genetics with the goal of synthesizing the methodology. While "bioinformatics" (computer science and biostatistics with application to biotechnology) is an area of explosive growth, for the Department of Biostatistics to lead bioinformatics efforts at Johns Hopkins would require us to commit the majority of our faculty positions to that subspecialty. Hence, it was decided that the best course is for us to recruit one or two faculty members in that area and to partner with other groups in the University to ensure that the best possible statistical reasoning is applied to biotechnology problems. It was also recognized that universities are unlikely to be competitive with industry in the development of specific biotechnology tools. Rather, our role will be to shape statistical thinking and methods underpinning new informatics tools.

There was enthusiasm for a new initiative in the foundations of statistical inference. Given the recent dramatic advances by Richard Royall and colleagues on likelihood methods for inference, it was proposed that we create momentum for a "transition-to-evidence-based statistics" movement. To promote this transition, departmental faculty might create a working group and plan how best to proceed.

Research opportunities at the interface of observational and experimental studies were discussed at length. Following recent changes to the Declaration of Helsinki for clinical research, which makes placebo-controlled clinical trials more difficult to justify, alumnus Frank Hurley raised the need for causal modeling from clinical trials as well as observational studies. Current work by Dan Scharfstein, Constantine Frangakis, and others will become increasingly relevant.

A related opportunity, introduced by Ron Brookmeyer, is the need for quantifying uncertainty from sources beyond sampling variation. In most observational or experimental studies, there is uncertainty about key parameters derived from the sample being different from the intended population, from the measured health outcomes being imperfect measures of the intended variable, from informative missing data, and other sources. By what process might these multiple sources of uncertainty be quantified, given a single study or many studies? What novel designs might improve our ability to quantify uncertainty?

/mja

Table 1. Retreat Participants

Name	Departmental Affiliation
Karen Bandeen-Roche	Faculty member
Karl Broman	Faculty member
Ron Brookmeyer	Faculty member
Frank Curriero	Faculty member
Marie Diener-West	Faculty member
Francesca Dominici	Faculty member
Constatine Frangakis	Faculty member
Liz Garrett	Faculty member
Steve Goodman	Faculty member
Rafael Irizarry	Faculty member
Elizabeth Johnson	Faculty member
Aidan McDermott	Faculty member
Giovanni Parmigiani	Faculty member
Steve Piantadosi	Faculty member
Bill Rising	Faculty member
Richard Royall	Faculty member
Dan Scharfstein	Faculty member
Rick Thompson	Faculty member
Jim Tonascia	Faculty member
Mei-Cheng Wang	Faculty member
Scott Zeger	Chair/Faculty member
Mitch Gail	Adjunct faculty member
Joanne Katz	Joint faculty member
Larry Moulton	Joint faculty member
Steve Self	Adjunct faculty member
Tim Wyant	Alumnus
Natalie Blades	Student
Weimin Chen	Student
Reg Dunn	Student
Wes Eddings	Student
Michael Griswold	Student
Fang-Chi Hsu	Student
Chiung-Yu Huang	Student
Hormuzd Katki	Student
Dongmei Liu	Student
Howard Mackey	Postdoctoral Fellow
Supreet Rangi	Student
Michelle Shardell	Student
Zhijin Wu	Student
Mary Joy Argo	Staff
Chris McCullough	Staff
Debra Moffitt	Staff
Kathy Spinnato	Staff

Table 2: Summary of Strengths and Weaknesses Identified by Retreat Participants

Strengths	Weaknesses
• Quality and diversity of faculty, staff, and students	• Information systems
• Commitment to statistical science: foundations ↔ methodology ↔ applications	• Disaggregation of statistical experts around the University and the difficulty of acting in concert, given a decentralized approach to administration
• Commitment to individual and collective excellent in research, education, and practice	• Inadequate expertise at the interface of computer science and statistics
• Environment (aka, 25¢ coffee): collegial; supportive competition; support for intellectual risk-taking	• Growth of teaching of biostatistics in departments other than Biostatistics
• Johns Hopkins Health Institutions: outstanding biomedical and public health colleagues who raise interesting biostatistical problems and help create/disseminate solutions	• The perception by colleagues in other departments that our introductory courses are too large or could be taught by statistically-oriented faculty outside Biostatistics
• Effective informal collaborations among primary and jointly-appointed faculty	• ScM program – non-competitive for price with other programs at state universities
• Opportunity for junior faculty to focus on establishing research careers	• Our ability to disseminate research results: e.g., over the web
• Public health and biomedical students taught by us become leaders around the world, maximizing our influence	• Ability to create attractive software that might become more widely used
• Decentralization, allowing us to create resources needed to compete with other top departments	• Student space
• Being a department in a school of public health, where biostatistics is seen as central to the school's mission, particularly by the current dean	
• Seminar program	
• Departmental website	

Table 3: Opportunities for Biostatistical Research

1. Non-invasive measurement systems
   - Biomonitoring: time series on many persons
   - Imaging
   - Gene expression arrays
   - Proteomics
   - Complex questionnaires – latent variable models

2. Statistical genetics/genomics/bioinformatics

3. Transition to evidence-based statistics

4. Integrated analyses of data from longitudinal studies: joint analyses of multivariate
   - Repeated measures
   - Times-to-events

5. Causal inference from observational and experimental studies
   - Statistical models with bias terms
   - Quantifying uncertainty beyond sampling variation

6. Clinical trials
   - Summarizing evidence
   - Measuring treatment effects with: partial compliance; drop-outs; treatment efficacy for subgroups
   - Combining evidence from many trials

7. Early detection of disease processes – biomarkers

8. Environmental epidemiology
   - Time and space risk models
   - Measurement error

9. Quality assurance for laboratory research
   - Pooling data
   - Variance components models

10. Teaching statistical reasoning and methods to health scientists and professionals

11. Internet-based data collection, management, measurement and analysis

12. Geographic Information Systems (GIS) for exploratory data analysis

13. Auditing/standards for professional practice

14. Quantifying health effects of increasing disparity of wealth: US urban health, developing vs. non-developing countries

15. Post-marketing surveillance of drug treatments

16. Vaccine development in international health research