Abstract word count: 128

Manuscript word count: 2964

Strategies for Developing Biostatistics Resources in an Academic Health Center

Leah J. Welty, Rickey E. Carter, Dianne Finkelstein, Frank E. Harrell Jr., Christopher J. Lindsell, Maurizio Macaluso, Madhu Mazumdar, Paul J. Nietert, Robert Oster, Brad H. Pollock, Paula K. Roberson, and James H. Ware

Leah J. Welty
Department of Preventive Medicine
Northwestern University Feinberg School of Medicine, Chicago, IL

Rickey E. Carter
Department of Health Sciences Research
Mayo Clinic, Rochester, MN

Dianne M. Finkelstein
Massachusetts General Hospital and Harvard University, Boston, MA

Frank E. Harrell Jr.
Department of Biostatistics
Vanderbilt University School of Medicine, Nashville, TN

Christopher J. Lindsell
Center for Clinical and Translational Science and Training
University of Cincinnati, Cincinnati, OH

Maurizio Macaluso
Division of Biostatistics and Epidemiology
Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Madhu Mazumdar
Division of Biostatistics and Epidemiology
Department of Public Health, Weill Cornell Medical College, New York, NY

Paul J. Nietert
Division of Biostatistics and Epidemiology
Department of Medicine, Medical University of South Carolina, Charleston, SC

Robert Oster
Division of Preventive Medicine, Department of Medicine
University of Alabama at Birmingham, Birmingham, AL

Brad H. Pollock
Department of Epidemiology and Biostatistics
University of Texas Health Science Center at San Antonio, San Antonio, TX

Paula K. Roberson
Department of Biostatistics
University of Arkansas for Medical Sciences, Little Rock, AR

James H. Ware
Department of Biostatistics
Harvard School of Public Health, Boston, MA

Acknowledgements

We thank all the members of the Biostatistics, Epidemiology, and Research Design (BERD) Key Function Committee of the Clinical and Translational Science Awards (CTSA) Consortium. We also thank the National Institutes of Health, and in particular, Laura Lee Johnson, PhD, of the National Center for Complementary and Alternative Medicine (NCCAM) and Iris Obrams, MD, MPH, PhD, of the National Center for Advancing Translational Sciences (NCATS).

This work was supported in part with federal funds from the NIH through the CTSA Program. The NIH CTSA funding was awarded to Northwestern University Feinberg School of Medicine (UL1 RR025741), Mayo Clinic (UL1 RR024150), the Harvard University School of Medicine (UL1 RR025758), Vanderbilt University (5UL1RR024975), the University of Cincinnati (UL1 RR026314), the Weill Cornell Medical College (UL1-RR024996), the Medical University of South Carolina (UL1RR029882), the University of Alabama at Birmingham (UL1RR025777), the University of Texas Health Science Center at San Antonio (UL1RR025767), and the University of Arkansas for Medical Sciences (UL1RR029884).

Abstract

Although biostatistics plays an important role in health-related research, biostatistics resources are often fragmented or ad hoc within Academic Health Centers (AHCs). Given the increasing complexity and quantity of health-related data, the emphasis on accelerating clinical and translational science, and the importance of reproducible research, there is interest in and need for the thoughtful development of biostatistics resources within ACHs. We have accordingly identified strategies for developing biostatistics resources in three areas: (1) recruiting and retaining biostatistics; (2) using biostatistics resources efficiently; and (3) improving science through biostatistics. Ultimately, we recommend AHCs centralize biostatistics resources in a unit rather than disperse them across clinical departments, as the former offers distinct advantages to investigator collaborators, biostatisticians, and ultimately to the success of the research and education missions of AHCs.

Introduction

Science is built upon rigorous experimentation and research. Biostatistics—the application of statistics to health-related research—provides powerful tools for developing scientific questions, designing valid studies, refining measurements, and analyzing data. A biostatistician's unique contribution to a research team is founded on quantifying uncertainty in and generating sound inferences from data. Because of the increasing complexity and quantity of health-related data, the need for biostatistical expertise is expanding and evolving.

In spite of its fundamental role, biostatistics resources are not uniformly well-established in academic health centers (AHCs). Although many experienced investigators value collaborating with biostatisticians, biostatistics is sometimes regarded as ancillary technology rather than academic discipline. As a result, some AHCs have lone biostatisticians scattered throughout clinical departments. Although other AHCs have centralized biostatistics groups, those groups that focus primarily on consultation provide a valuable service but fail to capitalize on the contributions biostatistics can make to interdisciplinary research.

We therefore propose focused strategies for AHCs to develop and support biostatistics resources. We elaborate on strategies in three specific areas: recruiting and retaining biostatisticians (Section 1); efficient use of biostatistics resources (Section 2); and improving science through biostatistics (Section 3). Our recommendations are directed primarily to senior leadership in AHCs, but they also apply to investigators with large research programs and senior biostatisticians interested in building capacity.

Our overarching recommendation is that the proposed strategies are readily executed in an AHC that creates and supports a centralized academic unit with biostatistics as a primary focus. Such units may take the form of a center, division, or department; we adopt the generic term 'biostatistics unit'

throughout. However, our strategies are relevant to all AHCs, from those that already support a centralized unit to those with few or dispersed biostatistics resources. Our remarks also apply to other disciplines, such as bioinformatics, that play similar roles in interdisciplinary research.

1. Recruiting and Retaining Biostatisticians

There is currently a shortage of biostatisticians qualified to meet the needs of AHCs[1]: statisticians—especially those with strong communications skills—are in high demand across a range of industries[2,3], and statisticians in AHCs must also be familiar with health-related research[4]. Efforts to attract and retain well-qualified biostatisticians should be a priority.

a. *Identifying Qualified Biostatisticians*

It is advantageous to have a senior biostatistician direct recruitment of new biostatisticians. Senior biostatisticians are uniquely positioned to help define the skills, experience, and training necessary to address research needs. For example, a clinical department's self-perceived need for a full-time doctoral-level biostatistician may be better addressed by a statistical programmer, with oversight from a (less than FTE) doctoral-level biostatistician. Senior biostatisticians have professional networks to identify candidates; candidates are more likely to respond to position announcements distributed by biostatistics units. When evaluating candidates, senior biostatisticians have disciplinary expertise not available to clinicians, other investigators, and many search firms. Just as most biostatisticians have limited ability to evaluate candidates for a clinical position, most clinicians lack the knowledge required to identify biostatisticians with appropriate technical background and experience.

For some investigators or department chairs, control over hiring outweighs the advantages of a senior biostatistician directing a search. In such cases, a respected partnership, which heavily involves

investigators and regards their candidate evaluations as critical, is preferred. Feedback from different perspectives – methodological training as well as collaborative potential – will result in higher quality, well-rounded hires with better chances for retention.

b. Creating an Attractive Environment

For most biostatisticians, opportunities for academic advancement are significant considerations when selecting a position. Too often, however, biostatisticians hired by principal investigators or clinical departments are expected to fill a perceived service role. Mismatches in expectations will negatively affect hiring and ultimately, retention. AHCs will find it easier to recruit and retain biostatisticians to an academic home that has a strong identity, spurs intellectual growth, values contributions to health-related research, and offers a clear and appropriate path for promotion. These attributes are readily fostered within a biostatistics unit where the biostatistician may collaborate with clinical investigators while working alongside other biostatisticians. For a junior biostatistician, it is essential to have senior biostatisticians to offer advice on effective collaboration and professional development. Biostatisticians, particularly those building careers, recognize that isolation from mentorship by other biostatisticians—as occurs when a sole biostatistician is hired within a clinical department—may limit opportunities for professional advancement.

To retain biostatisticians, promotion criteria must be aligned with the independent contributions biostatisticians make to interdisciplinary research. If being promoted requires external funding as principal investigator and a substantial portfolio of senior-author publications, an AHC cannot expect to retain biostatisticians whose primary focus is collaboration. Recently, several AHCs have redefined promotion criteria to better recognize the substantial intellectual contributions co-investigators make to research (e.g. [5,6]). Ultimately, AHCs with biostatistics departments may be best

able to retain biostatisticians because they have (1) the responsibility and resources to provide appropriate career development and (2) influence on appointments and promotions.

Biostatisticians will gravitate to a unit with strong leadership. Ideally, leaders should possess broad biostatistics knowledge, an impressive record of methodological and collaborative research, and exceptional communication skills. The leader's experience promotes health-related research, contributes to the unit's reputation, and sets an example for junior colleagues. Given the variety of disciplines the leader will encounter, he or she must also enjoy collaboration and possess generalist's skills. We strongly caution AHCs against recruiting junior faculty to start biostatistics units. Not only do junior biostatisticians lack the necessary research and administrative experience, but as the first member of a unit, they also want for guidance from senior biostatisticians and have limited ability to mentor new faculty recruits. Premature assumption of a leadership role may negatively affect the growth of the biostatistics unit in addition to the individual career of the junior biostatistician.

2. Building Efficiency

New technologies in areas as diverse as imaging, genetic sequencing, and electronic health records are fast outpacing the development and dissemination of applicable statistical methods[7]. The need for efficient use of biostatistics resources is rapidly increasing. Several strategies can maximize appropriate and effective utilization.

a. Specialization

Just as physicians specialize, most biostatisticians focus in one or two areas of statistics. The skills and knowledge pertaining to adaptive clinical trials, predictive modeling, and comparative-effectiveness research all differ. Doctoral programs in biostatistics cover the most widely used methods,

and any well-trained biostatistician can support entry-level health-related research. However, complex projects require specialized expertise. Clinical trials biostatisticians are able to design studies with the minimum sample size required to answer the clinical question; those with expertise in analyzing imaging data are familiar with methods to effectively extract signal from noise. Biostatistics units comprised of complementary specialists can accommodate disparate needs and will be well positioned to contribute to a variety of current and future health-related problems.

With a biostatistics unit, an investigator collaborating with a single biostatistician benefits from the collective expertise of the entire unit. Even when a biostatistician has methodological interests that align with an area of health-related research, the research will generate questions outside the biostatistician's expertise. In a centralized unit, biostatisticians may find the answer down the hall; when biostatisticians are isolated, they frequently reinvent the methodologic wheel – a process both inefficient and error prone. Furthermore, the leader of a centralized biostatistics unit can recognize redundant efforts and ensure that biostatisticians capitalize on collective expertise.

Biostatisticians also work more effectively by specializing in one or two medical areas. This occurs naturally when a biostatistician is well-integrated in a research team. For example, a biostatistician focusing on a particular disease learns the disease characteristics, treatment types, efficacy measures, and expected complications. Building this knowledge requires patience and commitment from both biostatistician and collaborators, but it is essential for optimal study design and data analysis. Collaboration is more successful and efficient when biostatisticians understand the substantive vocabulary and principles. In these cases, biostatisticians may also become scientific leaders within a medical area. A biostatistician's investment in the subject matter may ultimately matter more to the success of the collaboration than other factors, such as physical location.

b. Appropriate use of doctoral and master's level skills

Efficiency requires recognizing the training required to engage in different activities such as study design, data management, statistical programming, data analysis, manuscript preparation, and proposal development. A doctoral-level biostatistician necessarily possesses data management and programming skills, but so do master's-level biostatisticians, who are sometimes more proficient and often less expensive. If isolated doctoral-level biostatisticians are responsible for all data-related matters, they will engage disproportionately in data management and programming, spending less time on study design, statistical analysis or proposal development—areas where their expertise is pivotal. Many successful biostatistics units maintain a ratio of doctoral-level to master's-level personnel of about 1:2 or 1:3. If a clinical unit employs a single biostatistician, however, it may be less efficient to employ a master's-level rather than a doctoral-level biostatistician: (1) doctoral-level biostatisticians are more likely to have expertise in study design; (2) a grant proposal with a doctoral-level biostatistician as co-investigator will be reviewed more favorably than one with a master's-level biostatistician; and (3) although some master's-level statisticians function independently, most require oversight from a doctoral-level biostatistician. Doctoral-level biostatisticians should work independently as well as manage and mentor master's-level colleagues.

c. Infrastructure

Infrastructure also influences the efficiency of biostatistical collaboration. Biostatisticians require software and hardware that can be expensive and extensive. Software licenses and high-speed computing are not typically available except within specialized units and costs are rarely recouped in grant budgets. These resources are less expensive on a per-user basis when maintained centrally (e.g. site licenses, shared computing clusters). Once a biostatistics group reaches a critical mass of 5-10, it

also needs a system administrator to provide technical support and an administrator to manage grants, scheduling, and recruitments.

d. Centralization

With a centralized biostatistics unit, investigators requiring some biostatistics support may subscribe to a portion of a biostatistician's time. The research team does not need to commit to a full-time position, and biostatistics expertise is available to investigators who would otherwise go without. Having a critical mass of biostatisticians within a unit provides greater protection for transitions in employment and changes in funding levels; investigators have fewer concerns about continuity of funding or recruiting and retaining biostatisticians. A central biostatistics unit with dedicated institutional support can manage: (1) fluctuations in external support; (2) recruitment of new biostatisticians in anticipation of future needs; (3) and transitioning newly hired biostatisticians to a research team through training under senior leadership[8-10].

AHCs can also fund centralized biostatistics units to broadly support research and education. Biostatisticians who are not supported solely by external funding or clinical departments may participate in institutional activities, unfunded but promising collaborative projects, or teaching activities. For example, biostatisticians may serve on IRBs, mentor career development grant awardees, or give workshops.

A centralized biostatistics unit should balance sharing biostatistics resources with stretching them too thin. Although it may be inefficient for only a few groups within an AHC to have access to biostatisticians, it is equally inefficient for biostatisticians to divide their time among too many different projects. In the extreme, excessive sharing of biostatistics resources will result in the tragedy of the commons[11]: over-subscribed biostatisticians unable to meaningfully contribute to any specific project

and expending most of their energy on switching costs. Departments or groups with dedicated biostatisticians are not immune to this problem, especially if they are large, have many trainees, or underestimate the time required to conduct careful analyses.

Formal and thoughtful creation of units can help balance over-subscription. Some AHCs institute a "10% rule," discouraging biostatisticians from collaborating on projects for less than 10% effort. A similar strategy is to have each biostatistician maintain one or two larger collaborations, and spend a limited portion of their time on smaller projects (e.g. no more than three projects with <= 10% effort each). Because there are many demands on biostatistics resources, AHCs should adopt policies that prioritize use of biostatistics resources.

Although a centralized biostatistics unit offers many advantages in terms of efficiency, biostatisticians and their senior leadership must be responsive to collaborating investigators. No investigator wants to feel that a biostatistician in whom they invest resources and rely on for expertise is not fully invested in a project because of their location or because they are not funded at 100%. Strategies for managing this balance include: (1) biostatisticians having secondary appointments in the department of their primary and long-term collaborators; (2) biostatisticians focusing on one or two areas of subject matter expertise; and (3) biostatisticians having adequate master's level support to quickly perform routine statistical tasks. Communication about project timelines and competing demands is also a critical skill for biostatisticians in an AHC.

3. Improved Contributions to Science

Biostatistics is a rapidly changing field. Using methods tailored specifically to a research question results in increased precision and power compared to using tools previously developed for alternate purposes.

a. Continuing education

A central biostatistics unit can provide continuing education in statistical methodology and practice that individual clinical departments cannot. Biostatistics units offer regular access to methods-focused seminars and journal clubs, and informal conversations with other biostatisticians—all of which broaden and strengthen biostatisticians' knowledge and value as collaborators. A central unit may also dedicate financial resources for attending conferences or short courses in emerging areas of biostatistics. For example, a priority for biostatisticians (and investigators) is practicing reproducible analysis, reporting, and research. Exposure to reproducible research tools from colleagues or courses enhances the reproducibility of analyses and results. Biostatisticians affiliated with a central unit remain current in professional standards and areas of research by learning from each other and participating in continuing education activities.

b. Professional activities

Research teams glean maximum benefit from biostatisticians with strong professional qualifications. It is therefore wise for AHCs to provide resources for biostatisticians to engage in professional activities including: (1) preparing their own manuscripts and grant proposals; (2) presenting at conferences; and (3) providing peer-review for journals and funding agencies. These activities all benefit a biostatistician's collaborators. Biostatisticians with first-authored and methodological publications will be considered stronger co-investigators on research proposals, especially if they have published previously with the principal investigator. Time allocated for teaching or administrative responsibilities should not diminish time allocated for maintaining a program of research.

The time reserved for methodologic research varies across institutions; some AHCs consider 20% effort the minimum. This time may be funded from institutional support, indirect costs, clinical revenues, consulting income, or grants that include biostatistical aims. When biostatisticians are fully

engaged as collaborators with their clinical partners, research and development are fundamental activities for both parties. Biostatisticians should develop statistical methodology relevant to the particular research, and investigators should include sufficient allocation within grant budgets to support methodological research. Investigators should be aware that a statistical methods paper typically takes longer to write than a clinical paper; however, biostatisticians must be accountable for the time and resources provided.

c. Collaboration versus consultation

Biostatisticians need to be true collaborators with vested interests in the success of research projects. Too often, doctoral-level biostatisticians are perceived as technicians, not independent contributors. Aspects of data management may be classified as a service, but much of the research process requires independent intellectual contribution. Interpretation of results is often nuanced; establishing a thorough data analysis plan requires more than identifying the type of dependent variable and a list of independent variables and covariates. Although any trained biostatistician can perform basic power or sample size calculations, only one immersed in the field of inquiry will be able to design an efficient study directly addressing specific aims. Doctoral-level biostatisticians have a unique and specialized skill set that makes them an integral part of the research team. Their contributions are typically independent and implemented without supervision. Regardless of the structure of biostatistics resources, maximum scientific benefit is achieved when biostatisticians function as collaborators. Limiting biostatisticians to the role of consultant or service technician undermines their capacity to strengthen the foundations of scientific research.

d. Biostatistical literacy

Effective communication between investigators and biostatisticians is accomplished only in part by increasing the biostatisticians' familiarity with the appropriate medical field; investigators must also be familiar with the principles and methods of biostatistics. They should be aware of the assumptions and pitfalls of methods used to generate scientific evidence—both in their own studies and from the literature. Biostatisticians within AHCs should ensure that education for investigators adheres to high standards. Such education is continuous and occurs at many junctures—from one-hour consultations to formal classes to multi-year collaborations. It is encouraging that many AHCs now support master's-level programs in clinical research. Biostatisticians necessarily play an important role in meeting the new educational needs of the health-related research community, but isolated biostatisticians have limited ability to contribute. Biostatistics units provide a natural home for educational programs in biostatistics and research design. Furthermore, biostatisticians in centralized units are well positioned to mentor students who can contribute to data management, statistical analysis, consultation, and methodological research as part of their training.

Conclusions

With growing emphasis on clinical and translational research and evidence-based medicine, integrated research teams are increasingly essential in AHCs. The unprecedented quantity and complexity of health-related data requires that biostatisticians be key collaborators, a role they may readily assume when part of a central biostatistics unit. We have outlined strategies for building and supporting such biostatistics resources within an AHC, including strategies for recruiting and retaining biostatisticians; efficiently using biostatistics resources; and maximizing biostatistical contributions to scientific collaboration. A centralized biostatistics unit—with strong biostatistics leadership—is an essential evolution, with distinct advantages compared to a more fragmented organization. Just as the demand for patient-centered, integrated health care has galvanized the integration of physicians from multiple

specialties, so too should the growing complexity and importance of health-related research galvanize the thoughtful and careful structuring of biostatistics units as key elements in a successful research enterprise.

References

1. DeMets DL, Stormo G, Boehnke M, Louis TA, Taylor J, Dixon D. Training of the next generation of biostatisticians: a call to action in the U.S. *Statistics in medicine.* Oct 30 2006;25(20):3415-3429.

2. Lohr S. For Today's Graduate, Just One Word: Statistics. *The New York Times.* August 5, 2009;A1.

3. US Department of Labor. Occupational Outlook Handbook, 2010-2011.

4. Begg MD, Vaughan RD. Are Biostatistics Students Prepared to Succeed in the Era of Interdisciplinary Science? (And How Will We Know?). *The American Statistician.* 2011;65(2):71-79.

5. Vanderbilt University School of Medicine. Guidelines for Evaluating Faculty for Appointments and Promotions. https://medschool.vanderbilt.edu/faculty/details-appointments-and-promotions#guidelines. Accessed February 3, 2011.

6. University of Iowa. Faculty Handbook. http://www.public-health.uiowa.edu/faculty-staff/faculty/handbook/pdf/Chapter%204%20-%201.3%20-%20BIO.pdf. Accessed February 3, 2012.

7. Califf RM, Ginsburg GS. Organizational improvements to enhance modern clinical epidemiology. *JAMA : the journal of the American Medical Association.* Nov 19 2008;300(19):2300-2302.

8. Zelen M. The training of biostatistical scientists. *Statistics in medicine.* Nov 2003;22(21):3427-3430.

9. Tobi H, Kuik DJ, Bezemer PD, Ket P. Towards a curriculum for the consultant biostatistician: identification of central disciplines. *Statistics in medicine.* Dec 2001;20(24):3921-3929.

10. Gehan EA. The training of statisticians for cooperative clinical trials: a working statistician's viewpoint. *Biometrics.* Dec 1980;36(4):699-706.

11. Hardin G. The Tragedy of the Commons. *Science.* Dec 1968;162(5364):1243-1248.