



Published in final edited form as:

Stat Med. 2011 October 15; 30(23): 2767–2777. doi:10.1002/sim.4184.

Evaluation Metrics for Biostatistical and Epidemiological Collaborations

Doris McGartland Rubio^{1,*†}, Deborah J. del Junco², Rafia Bhore³, Christopher J. Lindsell⁴, Robert A. Oster⁵, Knut M. Wittkowski⁶, Leah J. Welty⁷, Yi-Ju Li⁸, Dave DeMets⁹, and the Biostatistics, Epidemiology, and Research Design (BERD) Key Function Committee of the Clinical and Translational Science Awards (CTSA) Consortium

¹ Data Center, Center for Research on Health Care, Division of General Internal Medicine, Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A

² Biostatistics, Epidemiology, and Research Design Core, Center for Clinical and Translational Sciences, University of Texas Health Science Center, Houston, Texas, U.S.A

³ Division of Biostatistics, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas, U.S.A

⁴ Center for Clinical and Translational Science and Training and Department of Emergency Medicine, University of Cincinnati, Cincinnati, Ohio, U.S.A

⁵ Division of Preventive Medicine, Department of Medicine, and Design and Biostatistics Program, Center for Clinical and Translational Science, University of Alabama at Birmingham, Birmingham, Alabama, U.S.A

⁶ Biostatistics, Epidemiology, and Research Design Core, Center for Clinical and Translational Science, The Rockefeller University, New York, New York, U.S.A

⁷ Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, U.S.A

⁸ Department of Biostatistics and Bioinformatics, Duke Translational Medicine Institute, Center for Human Genetics, Duke University Medical Center, Durham, North Carolina, U.S.A

⁹ Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, Madison, Wisconsin, U.S.A

Abstract

Increasing demands for evidence-based medicine and for the translation of biomedical research into individual and public health benefit have been accompanied by the proliferation of special units that offer expertise in biostatistics, epidemiology, and research design (BERD) within academic health centers. Objective metrics that can be used to evaluate, track, and improve the performance of these BERD units are critical to their successful establishment and sustainable future. To develop a set of reliable but versatile metrics that can be adapted easily to different environments and evolving needs, we consulted with members of BERD units from the consortium of academic health centers funded by the Clinical and Translational Science Award Program of the National Institutes of Health. Through a systematic process of consensus building and document drafting, we formulated metrics that covered the three identified domains of BERD

*Correspondence to: Doris Rubio, Division of General Internal Medicine, Department of Medicine, School of Medicine, University of Pittsburgh, 200 Meyran Avenue, Suite 200, Pittsburgh, PA 15213, U.S.A.

†rubiodm@upmc.edu

practices: the development and maintenance of collaborations with clinical and translational science investigators, the application of BERD-related methods to clinical and translational research, and the discovery of novel BERD-related methodologies. In this article, we describe the set of metrics and advocate their use for evaluating BERD practices. The routine application, comparison of findings across diverse BERD units, and ongoing refinement of the metrics will identify trends, facilitate meaningful changes, and ultimately enhance the contribution of BERD activities to biomedical research.

Keywords

metrics; biomedical research; collaboration; clinical and translational science; biostatistics; epidemiology; research design

1. INTRODUCTION

Skills in biostatistics, epidemiology, and research design (BERD) are indispensable for the success of biomedical research. Coupled with ever-increasing computational capacity and biomedical knowledge, BERD methodologies have evolved to describe and model the various interactions between biology, medicine, genetics, society, the environment, and health. Yet additional advances in statistical methods are urgently needed to improve the ability of investigators in clinical and translational research to draw valid inferences about complex biological, epidemiological, and health systems. Because the understanding of modern scientific literature and the interpretation of study data require scientists and clinicians to have a firm grasp of the underlying principles of BERD, the core curricula of research training programs are focusing on these principles. This is being accompanied by a proliferation of units that provide BERD expertise within academic health centers.

BERD expertise and practices within academic institutions are organized in various types of structures (e.g., departments, divisions, and centers), and multiple structures often overlap within a single institution. BERD practices may be associated with public health or clinical research enterprises, and specialization in fields such as computational biostatistics or statistical genetics links BERD practices with bioinformatics. At some institutions, BERD is recognized as an academic discipline with practitioners engaged in teaching, engaged as collaborators, or as independent investigators. At other institutions, BERD practitioners are viewed as providers of support services. Not infrequently, both roles for BERD occur within institutions. This diversity is often reflected in research teams in which the role of the BERD practitioner can range from an ad hoc consultant to a key investigator. Regardless of how BERD activities are integrated into an academic health center, the assumption is that they will substantially contribute to our knowledge and understanding of biomedical research. The extent of their contribution, however, remains largely unquantified because of the lack of metrics for their evaluation.

Evaluating performance in interdisciplinary biomedical research is highly complex, requiring a pluralistic approach that extends beyond conventional metrics [1]. New approaches to evaluation link research activities to spatially and temporally remote outcomes through, for example, the systematic logic model of the National Institute of Environmental Health Sciences [2]. The framework of the logic model clearly differentiates between academic outputs (e.g., publications and grants) and outcomes such as changed health, changed practice, and changed behavior, although the proposed metrics primarily reflect the traditional outputs by tracking publications, presentations, and grants. Since traditional metrics offer only incomplete evaluation and there is a growing interest in evaluating the spectrum of the academic research enterprise, from administration to diverse

medical specialties [3–5], the focus of BERD practice on developing quantitative metrics for evaluation is both vital and timely.

Evaluation provides the vehicle to document the areas of strength and areas for improvement for the BERD unit as a whole so that academic health centers can develop a strong infrastructure in biostatistics and epidemiology. With the increasing demand for biostatistical and epidemiological expertise, BERD units need to understand how to leverage the various strengths of their BERD practitioners. Individually, evaluation benefits each BERD practitioner by providing direct information about his or her strengths and accomplishments as well as areas for improvement. Clinical and Translational Science is an exploding field. For practitioners to be effective collaborators, they must understand how to leverage their individual strengths and address their weaknesses so that they can maximize their skills. Evaluation provides critical information to BERD practitioners as to how their time and expertise is spent, enabling them to adjust their focus as appropriate.

The establishment of Clinical and Translational Science Awards (CTSAs) by the National Institutes of Health (NIH) and the provision of NIH funds for BERD units and other clinical and translational research infrastructure have brought the need for evaluation metrics to the forefront. Each BERD unit is required to annually report on their activities and accomplishments. At its inception in 2006, the BERD Key Function Committee recognized both the need for evaluating BERD activities and the need for developing a broad consensus about the evaluation metrics to be used. In this article, we begin by describing the approach that members of the committee used to achieve consensus on the metrics. We then provide definitions of broadly applicable metrics and describe the utility of these metrics for evaluating the contribution of BERD activities to academic health centers.

2. CONSENSUS DEVELOPMENT OF METRICS

The BERD Key Function Committee includes representatives from all CTSA-funded institutions and from the NIH. One of the first acts the Committee undertook was to form the BERD Evaluation Subcommittee to develop evaluation guidelines and metrics.

The subcommittee began by reviewing the evaluation metrics proposed in the set of CTSA applications that were originally funded. Of the 24 institutions awarded CTSAs during the first two review cycles, 19 (79%) provided summaries of their BERD activities. When the subcommittee reviewed the summaries, six categories of measurable BERD activities emerged: (1) consultations with clinical and translational science investigators; (2) grant applications submitted and funded; (3) protocols developed and reviewed; (4) abstracts and manuscripts submitted and accepted; (5) new methodologies developed, applied, and distributed; and (6) educational activities, courses, and students.

We formed subcommittees to develop metrics for each of these six categories. During monthly conference calls, the group leaders reported on progress and received feedback from subcommittee members. The subcommittee then began to collate the metrics into a working document. Collectively, it was agreed that the evaluation metrics document must be a “living” document that is able to accommodate the inherently transformational nature of activities occurring at CTSA-funded institutions. The document must also reflect the fact that academic health centers vary considerably in structure and strengths. As BERD functions vary and evolve, not all metrics will apply to all situations, and metrics will need to evolve.

The subcommittee presented a draft of the evaluation metrics document to the BERD Key Function Committee for comment and use by all of its members. Rather than initiating a formal Delphi process, the subcommittee posted the document on the CTSA BERD Wiki

site. Since 2008, the BERD Key Function Committee members have accessed the living document for review, comments, and suggested edits, and additions are reviewed monthly.

The committee and subcommittee members noted that while the NIH requires each CTSA-funded institution to report on numerous criteria and while the CTSA provides a forum for the development of metrics, the metrics should be broadly relevant to all BERD units within academic health centers, regardless of their funding mechanism. In addition, the metrics should be sufficiently flexible to allow for differentially evaluating subcomponents of a single BERD unit, such as might occur when there are multiple funding sources to which the unit is accountable. Members also noted that some activities may be performed by personnel other than those formally associated with a BERD unit or by personnel who are part of a BERD unit but do not specialize in BERD practices. Redundancies are often essential to the health of the BERD community overall, and overlap in functions and responsibilities can offer valuable opportunities for cross-component and cross-disciplinary synergies and efficiencies. We agree that the intent of the metrics is to provide the BERD community as a whole with evaluation measures that reflect the contributions that BERD activities make across academic health centers.

Since the first living metrics document was made available to CTSA members, a significant maturation has occurred. As new members join the BERD Key Function Committee, understanding of how BERD activities integrate with the clinical and translational sciences has crystallized, and the current iteration of the metrics represents the consensus experience of BERD practitioners at 46 institutions across the United States. By making these metrics publicly available, we are providing a tool to help justify the allocation of resources for BERD activities and to facilitate the sound management of these resources. We advocate that by using these metrics, BERD units can systematically assess their impact on individual academic health centers as well as on clinical and translational research as a whole.

3. EVALUATION FRAMEWORK

The six categories of measurable BERD activities originally identified have been encapsulated into the following three domains:

1. Development and maintenance of collaborations with clinical and translational science investigators,
2. Application of BERD-related methods to clinical and translational research, and
3. Discovery of novel BERD-related methodologies.

The three domains and their interrelationships are shown in Figure 1. As the figure indicates, evaluation metrics can apply either within a domain or at the intersection of domains. The essence of what BERD units should accomplish can be broadly stated as having an impact on biomedical research, and this is central to the evaluation framework. It should be noted that while overlapping, the domains are distinct and reflect differing focus areas. A BERD unit that is well integrated into an academic health center would be expected to score well in the first domain, while one that is successful in supporting investigators in grant submissions and publications would score well in the second domain. A BERD unit that maintains the capacity to evolve new methods, whether practical or theoretical, would be expected to perform well in the third domain. BERD units should consider evaluation within this framework as an essential component of quantifying their impact. In subsequent sections of this article, we discuss the metrics relevant to each of the three domains and to the intersection of domains so that BERD units can use these metrics to evaluate their efforts.

4. METRICS

4.1. Development and maintenance of collaborations with clinical and translational science investigators

For the first domain, we focused on consultations, collaborations (intellectual partnerships between BERD practitioners and investigators), investigator education, investigator mentoring, and investigator satisfaction. These activities should be tracked objectively and lend themselves to the metrics of effort or time applied, the number of individuals taught and mentored, and other measures outlined in Table I.

Special attention should be paid to the tracking of effort since BERD activities can span all three domains. Often, a BERD practitioner needs to spend time researching established methods or developing a new methodology for a project. Interesting concepts may arise during the course of the research so that a manuscript on a method appropriately reflecting these concepts is warranted. Whether and how such time is kept track of and accounted for remains a matter of debate. For example, if the development of a new method arises as a result of a consultation, should the hours spent developing the method or writing the manuscript be attributed to and funded by the project? What if the method turns out to be inappropriate for the particular project? Are the hours directly or even indirectly attributable to the consultation? Furthermore, how accurately can and should the hours in these cases be counted? In the absence of rigorous accounting methods, effort should at least be monitored in terms of ordinal categories, such as small, medium, and large effort, with a project requiring less than 10 hours considered to be a small effort and a project requiring more than 50 hours considered to be a large effort. While the choice between detailed or categorical tracking will depend on the needs of the BERD unit and the BERD practitioner, some measure of the effort applied to these activities should be considered.

Tracking the mentoring of investigators, including junior biostatisticians and epidemiologists, is particularly important, although the metrics for this are also challenging. For example, the difference between investigator education and investigator mentoring can be subtle [6]. Metrics for mentoring first require an understanding of what constitutes mentoring in general, followed by an understanding of the specific role of BERD practitioners in mentoring clinical and translational science investigators. While it is possible to track time, effort, number, and type of mentoring activities, the as-yet-undefined nature of outcomes from successful mentoring of investigators by BERD practitioners lends itself to a qualitative approach, such as through the use of success stories. BERD units should capture these success stories as models of their contribution.

At their core, BERD activities require teamwork and interaction between an investigator and a BERD practitioner. In this regard, a component of success is the satisfaction of both the investigator and the practitioner. Tracking appropriate aspects of satisfaction has great utility in the quality feedback loop, and should be incorporated into the evaluation of a BERD unit. We have characterized satisfaction aspects in terms of perceptions of timeliness, professionalism, collegiality, efficiency, and the knowledge/skill base. Satisfaction should always be considered relative to the type of service provided (consultation, collaboration, education, or mentoring) and the possibility of personal biases.

4.2. Application of BERD-related methods to clinical and translational research

For the second domain, we developed metrics regarding proposals, grants, and contracts; dissemination of results in abstracts, presentations, and publication of manuscripts; and scientific integrity review and oversight of research. These metrics are outlined in Table II.

The application of BERD-related methods to clinical and translational research should result in the implementation of new studies. One of the critical contributions of BERD practitioners to the success of clinical and translational science investigators, especially new investigators, is to help them obtain funding for their research by advising them on sound research design methods and assisting them with the development of their grant proposals. BERD units should track the number of proposals submitted, resubmitted, and funded in order to demonstrate the value added by BERD practitioners. Moreover, the receipt of funding provides justification for BERD practitioners to recover the cost of their efforts.

In the United States, tracking BERD efforts related to assistance with grant writing is complicated by the current policy of the Office of Management and Budget. According to this policy, federally funded BERD practitioners can provide advice on study design, including advice about sample size, randomization and sampling plans, monitoring plans, and analysis plans in general, but they are formally prohibited from receiving direct grant funds for participating in the preparation and assembly of the application for federal funding. This restriction applies even if the federal funding is provided as support for the BERD unit to undertake consultation services. To ensure available time for proposal development, BERD units should receive some funding from nonfederal sources, such as institutional funds or from indirect funds.

During the active phase of a project, BERD practitioners may be involved in a variety of consulting and collaborative activities (e.g., designing data collection tools and survey instruments, implementing sampling and randomization schemes, monitoring data quality, monitoring safety, and performing interim analyses), and the time they spend on these activities is addressed in Table I. However, both during and after a project, they are also involved in dissemination of research findings, typically through abstracts, presentations, and manuscripts, and the metrics for tracking these activities are addressed in Table II. Some widely recognized publication metrics, such as journal impact factors and citation analyses, are frequently misused, and their limitations should be recognized [8,9]. Other publication indices, such as the *h*-index [Hirsch 2005], a single metric representing both the number of articles published and the impact of the articles in terms of citations, may facilitate comparisons across institutions [7]. Although BERD practitioners should receive consistent and appropriate credit for contributions, it is important to remember that the inclusion of a BERD practitioner as a coauthor is not the only criterion for identifying involvement. If an abstract or article acknowledges BERD support, this should be considered evidence of BERD involvement. This information should be documented so that the accurate involvement of BERD is reflected.

In addition to participating in the writing of grant proposals, BERD practitioners engage in professional and institutional service activities that promote scientific credibility through the appropriate application of BERD methods. Peer-reviewing manuscripts, serving as a statistical reviewer, or participating on an editorial board are all activities essential to ensuring the validity of research findings. Either for their own institutions or for external agencies, BERD practitioners review grant proposals for funding decisions, for the protection of human subjects, and for scientific or methodological acceptability. They are also involved in study oversight through their participation in Data Monitoring Committees, institutional review boards, ethics committees, and advisory committees. Other professional service activities, such as organizing conferences or leadership in professional societies, also play a role in enhancing clinical and translational research. Table II includes metrics for evaluating professional and institutional service activities that should be used to quantify these activities.

4.3. Discovery of novel BERD-related methodologies

While many statistical methods are available for clinical research, there remains much to be done to develop methods for translational science, i.e., for assessing the impact of environmental factors, the severity of complex phenotypes, and the effectiveness of intervention strategies in real-world settings [5,10,11], including complex populations with multiple morbidities and diversity of personal or cultural preferences. For today's clinical and translational investigations (e.g., "omics" research, personalized diagnostics and interventions as well as comparative effectiveness research with small or highly stratified patient populations), flexible design and computationally efficient methods for analyzing large and complex datasets are increasingly sought after commodities. The emerging enormous data warehouses from gene expression and genotyping studies and integration across multiple electronic medical record systems rapidly increase the demand for computational biostatistics resources. These vast and complicated data sets present unique challenges for conventional statistical methods developed to address more narrowly focused research hypotheses under more controlled research conditions.

With challenges come opportunities, and it is important to track a BERD unit's contribution to developing and implementing appropriate methods. Nonparametric statistical methods are well suited for analyzing complex genetic [12], environmental, and community settings [13], while Bayesian methods are particularly well suited for assessing changes in beliefs (e.g., changes in patients' preferences regarding risks and benefits of interventions) and for optimizing clinical trial design. However, many of these approaches are computationally more demanding than traditional methods, and the validity of the bioinformatics implementations may vary [14,15]. Fortunately, recent advances in computational biostatistics have overcome many hurdles, opening the door for BERD practitioners to develop new statistical methods and computational techniques that are better suited for the emerging needs of clinical and translational researchers. These needs include the development of new and novel approaches to the sampling, recruitment, and retention of study participants; the design of scales and instruments; the monitoring of clinical trials; the evaluation of research enterprises and phenotype-genomics relationships; and the strengthening of the interface between BERD and biomedical informatics.

To provide clinical and translational scientists with the computational biostatistics tools they need, BERD practitioners should be spearheading the development of new methods, conducting studies to assess the validity of the new methods, and publishing the results of both the methodological advances and their impact on applications. Dissemination of new methods by developing Web tools, software packages, and educational materials and by reporting on elegant solutions to complex practical and theoretical study design questions will extend the impact of BERD on clinical and translational research. Table III lists metrics for evaluating the discovery of new BERD-related methodology that BERD units should use to assess their contributions in this domain

5. DERIVING HIGHER-LEVEL METRICS

After each metric has been tracked, the problem arises of how to aggregate the data into meaningful scores. Simply scoring each metric individually and then taking the sum of the scores implicitly assigns the same relative weight to each activity and will be unhelpful or misleading. For example, adding the number of Web sites developed to the number of grants awarded does not appropriately reflect the relative benefit and impact of the efforts. This problem is not unique to evaluation and is encountered in the clinical and translational research itself, for example, when numerous phenotype markers need to be considered to assess the severity of disease and each marker has a different effect or impact. Methodological research to advance clinical and translational science by providing better

statistical methods for dealing with multivariate environmental, genetic, genomic, and phenotypic variables may also help with developing better metrics for evaluating the impact of BERD on clinical and translational research. As new methods are developed and the field of clinical and translational sciences advances, we need to re-evaluate, refine and add metrics so that we can continue to adequately evaluate BERD units and their contribution to the field.

The importance of the three evaluation domains will vary according to a BERD unit's mission. Some BERD units have a service focus while others might emphasize methodological research or teaching. We recommend assessing metrics in the three domains separately. Just as combining metrics within a domain can be misleading, a composite score necessarily places equal importance on each domain.

6. APPLYING THE METRICS

There are many ways in which an institution can obtain and use evaluation metrics. Although most BERD units will collect similar information as other units within their institution or BERD units in other institutions, they may choose to customize the metrics to meet their individual evaluation needs. For example, details of protocols submitted and funded, abstracts and manuscripts submitted and accepted, and educational activities may well be incorporated in institutional tracking systems as the traditional metrics of academic success. For BERD practitioners, investigator satisfaction and effort tracking on individual projects are relatively new metrics that may take time and a culture change to fully integrate into daily activities. Adding such information to the electronic tracking systems that are becoming increasingly common will make this information readily accessible and reportable. Doing so in real time will ensure that evaluation in terms of the institution's selected metrics leads to ongoing improvements in the procedures and structure of the BERD unit, especially as applied to meeting the needs of clinical and translational science investigators.

Not all of the investigators' needs are represented in the metrics described here, so additional metrics will need to be added to keep pace with new directions and knowledge in the field. Moreover, depending on infrastructure, a BERD unit may need to modify the metrics to ensure that evaluation is feasible. Given the evolving environment for clinical and translational research and the rapidly changing pace of infrastructure development, the tracking systems used to collect data for evaluation metrics should be easily modifiable. While flexibility in the details of the metrics is essential to their broad implementation, each BERD unit should use the three-domain framework and definitions provided here.

By evaluating BERD activities consistently within this common framework, the information needed for strategic development of BERD units in academic health centers will become more readily available. Within an institution, identifying voids in BERD skills or functions might drive a change in hiring or training practices; lack of timely responses to the investigators' requests for consultations might indicate the need to find efficiencies or to obtain additional resources; and identifying particular strengths might encourage new and fruitful areas of biomedical and BERD research. Data demonstrating the numerous ways that BERD practitioners contribute to protocol development and implementation, abstract and manuscript preparation, methodology development, educational activities, and mentoring of junior investigators will ultimately allow the staff and administrators of BERD units to demonstrate the impact of the units on the overall success of academic health centers. Communicating and disseminating the results of evaluations should lead to increased recognition of the contributions that BERD units make to clinical and translational research and should also garner funding and support for these contributions.

The BERD Evaluation Subcommittee has begun to systematically collect and analyze the evaluation metrics in a voluntary survey of CTSA-funded institutions. Initially, the group collected data that characterize the organizational structure and function of participating BERD units. These data will allow results from the evaluation metrics survey to be standardized. Compiling the standardized results will provide a baseline distribution describing the performance of CTSA-funded BERD units using the major indices discussed here, and data from annual surveys will enable time trend assessments. The findings are expected to stimulate dialogue and facilitate the development of consensus-based benchmarks for best BERD practices that are informed by data, and they will pave the way for continuing quality improvement and long-term strategic planning.

7. CONCLUSION

The consensus process for establishing evaluation metrics has identified 56 performance items in three domains that should be routinely assessed with quantitative measures that are typically available to a broad array of BERD units, regardless of their structure and function. The areas of performance are related to the development and maintenance of collaborations with clinical and translational science investigators (Table I), the application of BERD-related methods to clinical and translational research (Table II), and the discovery of novel BERD-related methodologies (Table III). We realize that all academic health centers are not alike. However, the key metrics reported here are generalizable to most if not all of them and have been defined to have sufficient reliability across institutions. The metrics should be used to provide an evidence base for changing and improving BERD practices in biomedical research. Additional metrics will continue to be developed and will focus on the more complex and more abstract domains.

The BERD Evaluation Subcommittee's living document, a comprehensive set of evaluation guidelines, is currently available to the public via an open-source Web site, CTSpedia [16]. To ensure maximal applicability and global benefit of these guidelines, development will continue in this open forum for all BERD practitioners working with clinical and translational researchers, regardless of affiliation or funding source. Refinements to the living document will occur as we gain more experience in applying the evaluation metrics, comparing results across BERD units, and identifying important new domains of BERD performance amenable to measurement.

Measuring the impact of BERD activities on biomedical research as a whole is the ultimate goal, and important future directions will include documenting success stories, tracking the mentoring of investigators, and assessing BERD contributions to national and international research networks and organizations. As metrics to evaluate the full spectrum of BERD's influence on biomedical research continue to evolve, the data will enable us to venture beyond justifying and managing resources at a single BERD unit to setting the course for the sustained growth and direction of BERD practices.

Acknowledgments

We thank all of the members of the Biostatistics, Epidemiology, and Research Design (BERD) Key Function Committee of the Clinical and Translational Science Award (CTSA) Program and especially the members of the BERD Evaluation Subcommittee. We are particularly grateful to the following individuals, who provided substantial commentary on this manuscript: Shelley Hurwitz, PhD, Harvard Medical School; Judith D. Goldberg, ScD, New York University School of Medicine; David Jarjoura, PhD, Ohio State University; Sally W. Thurston, PhD, University of Rochester; Gerald Beck, PhD, Cleveland Clinic; Brad H. Pollock, PhD, University of Texas Health Science Center, San Antonio; Rickey E. Carter, PhD, Mayo Clinic; Paul J. Nietert, PhD, Medical University of South Carolina; Peter Peduzzi, PhD, Yale School of Public Health; James Dziura, PhD, Yale School of Medicine; and Mary Lindstrom, PhD, University of Wisconsin. We also thank the National Institutes of Health

(NIH) and, in particular, Iris Abrams, MD, MPH, PhD, of the National Center for Research Resources (NCRR) and Dennis Dixon, PhD, of the National Institute of Allergy and Infectious Diseases.

This article was approved by the CTSA Consortium Publications Committee. The project reported here was supported in part with federal funds from the NCRR and NIH through the CTSA Program, which is a component of the Roadmap Initiative, Re-Engineering the Clinical Research Enterprise. The NCRR/NIH CTSA funding was awarded to the University of Pittsburgh (UL1 RR024153), the University of Texas Health Science Center at Houston (UL1 RR024148), the University of Texas Southwestern Medical Center at Dallas (UL1 RR024982), the University of Cincinnati (UL1 RR026314), the University of Alabama at Birmingham (UL1 RR025777), Rockefeller University (UL1 RR024143), Northwestern University Feinberg School of Medicine (UL1 RR025741), Duke University Medical Center (UL1 RR024128), and the University of Wisconsin at Madison (UL1 RR025011).

References

1. Klein JT. Evaluation of interdisciplinary and transdisciplinary research: a literature review. *American Journal of Preventive Medicine*. 2008; 35 (2 Suppl):S116–123. [PubMed: 18619391]
2. Engel-Cox JA, Van Houten B, Phelps J, Rose SW. Conceptual model of comprehensive research metrics for improved human health and environment. *Environmental Health Perspectives*. 2008; 116:583–592. [PubMed: 18470312]
3. Nolte KB, Stewart DM, O’Hair KC, Gannon WL, Briggs MS, Barron AM, Pointer J, Larson RS. Speaking the right language: the scientific method as a framework for a continuous quality improvement program within academic medical research compliance units. *Academic Medicine*. 2008; 83:941–948. [PubMed: 18820524]
4. Heinemann AW. Metrics of rehabilitation research capacity. *American Journal of Physical Medicine and Rehabilitation*. 2005; 84:1009–1019. [PubMed: 16327420]
5. Baren JM, Middleton MK, Kaji AH, O’Connor RE, Lindsell C, Weik TS, Lewis RJ. Evaluating emergency care research networks: what are the right metrics? *Academic Emergency Medicine*. 2009; 16:1010–1013. [PubMed: 19732037]
6. Deutsch R, Hurwitz S, Janosky J, Oster R. The role of education in biostatistical consulting. *Statistics in Medicine*. 2007; 26:709–720. [PubMed: 16625526]
7. Thompson DF, Callen EC, Nahata MC. New indices in scholarship assessment. *American Journal of Pharmaceutical Education*. 2009; 73 article 111.
8. Pendlebury DA. The use and misuse of journal metrics and other citation indicators. *Archivum immunologiae et therapiae experimentalis*. 2009; 57:1–11. [PubMed: 19219526]
9. Moed HF. New developments in the use of citation analysis in research evaluation. *Archivum immunologiae et therapiae experimentalis*. 2009; 57:13–18. [PubMed: 19219533]
10. DeMets DL. Statistical issues arising in the Women’s Health Initiative: discussion. *Biometrics*. 2005; 61:914–918.
11. del Junco DJ, Vernon SW, Coan SP, Tiro JA, Bastian LA, Savas LS, Perz CA, Lairson DR, Chan W, Warrick C, McQueen A, Rakowski W. Promoting regular mammography screening. I. A systematic assessment of validity in a randomized trial. *Journal of the National Cancer Institute*. 2008; 100:333–346. [PubMed: 18314473]
12. Morales JF, Song T, Auerbach AD, Wittkowski KM. Phenotyping genetic diseases using an extension of mu-scores for multivariate data. *Statistical Applications in Genetics and Molecular Biology*. 2008; 7 article 19.
13. Diana M, Song TT, Wittkowski KM. Studying travel-related individual assessments and desires by combining hierarchically structured original variables. *Transportation*. 2009; 36:187–206. [PubMed: 20953273]
14. Oster RA, Hilbe JM. An examination of statistical software packages for parametric and nonparametric data analyses using exact methods. *American Statistician*. 2008; 62:74–84.
15. Oster RA, Hilbe JM. Rejoinder to “An examination of statistical software packages for parametric and nonparametric data analyses using exact methods. *American Statistician*. 2008; 62:173–176.
16. CTSpedia, the knowledge base for clinical and translational research. [Accessed June 9, 2010.]. <http://www.ctspedia.org/do/view/BERDConsortia/BERDEval>

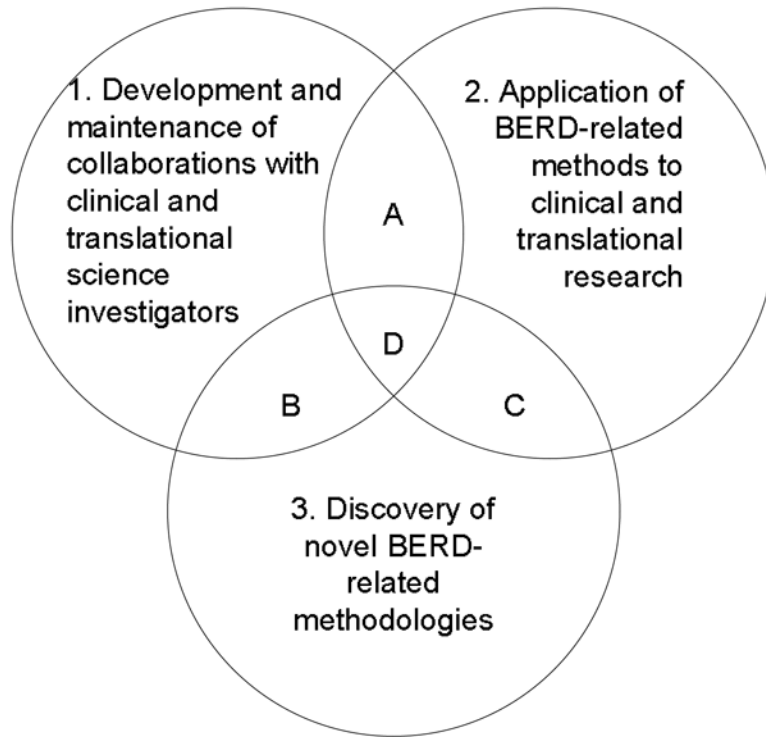


Figure 1.

The three domains of metrics to evaluate biostatistical, epidemiological, and research design (BERD) activities performed in academic health centers. Evaluation metrics can apply either within a domain (1, 2, or 3) or at the intersection of domains (A, B, C, and D). At the center of the evaluation framework (D) lies the intersection of the domains, an ideal opportunity for BERD units to maximize their impact on biomedical research.

Table I

Definitions and metrics for the development and maintenance of collaborations with clinical and translational science investigators

Definitions

- BERD practitioner—A specialist in biostatistics, epidemiology, or research design (BERD).
 - Clinical and translational science investigator An—individual who is pursuing biomedical research for the purpose of increasing knowledge that will contribute to the improvement of human health.
 - Collaboration—An intellectual partnership between a BERD practitioner and an investigator. Collaboration involves applying a multidimensional team approach to the scientific method.
 - Consultation—An interaction between an investigator and a BERD practitioner for the purpose of conferring about biostatistics, epidemiology, or research design. Consultations are generally limited in scope and are aimed at completing a requested activity for the investigator. A consultation may involve formal interactions, such as face-to-face meetings or workshops, or it may involve informal interactions, such as telephone calls, e-mail communications, or hallway discussions.
 - Education—The delivery of knowledge pertaining to biostatistics, epidemiology, or research design to investigators. Education may be formal or informal and may be didactic or use alternative means.
 - Investigator satisfaction—The investigator's perceptions about whether and how his or her needs have been met.
 - Mentoring—A developmental activity by which the BERD practitioner imparts knowledge and skills to the investigator, with the goal of fostering the investigator's professional growth.
-

Metrics

Consultation

- Time or effort spent on activities (hours or small/medium/large effort)
- Number of consultations
- Number of investigators

Collaboration

- Inclusion of a BERD practitioner in an activity (as co-investigator on a grant, as coauthor on a manuscript, as coauthor on an abstract, etc.)
- Number of study teams created

Education

- Type of activity (lecture, workshop, online class, etc.)
- Program affiliation (institution, college, department)
- Level of students
- Time applied to prepare, teach, and grade (credit hours, contact hours)
- Number of students
- Student grades
- Student course evaluations
- Peer evaluations

Mentoring

- Time or effort devoted to mentoring activities
- Number of investigators mentored
- Level of investigators mentored
- Types of mentoring activities
- Involvement in career development grants and awards
- Success stories

Definitions

Investigator satisfaction

- Timeliness
 - Professionalism
 - Collegiality
 - Efficiency
 - Knowledge/skill base
-

Table II

Definitions and metrics for the application of BERD-related methods to clinical and translational research *

Definitions

- Abstracts and presentations—Summaries of research findings presented at scientific meetings or similar public forums. Selection of abstracts for presentation may be by peer review, and the abstracts may be published in conference proceedings or journal supplements. Summaries of research first presented by the media are not classified as abstracts.
- BERD contribution—BERD practitioners may contribute to a wide range of research activities, including study design; aims and hypotheses development; sample size and power calculation; statistical analyses (plan, interim, and final); data collection and management forms and systems; data safety and monitoring plans; participant recruitment, sampling, and randomization methods; and authoring of reports.
- Contract—A legal agreement that typically binds an investigator to conducting specified research activities or services in return for payment.
- Funding source—Funding may be intramural (within the investigator's institution), extramural (outside the investigator's institution), or a combination of both. Extramural sources include private foundations, industries, and regional or national government agencies.
- Grant—Typically, a monetary award to support activities outlined in a proposal. A grant may be unrestricted if the allocation of funds is not prespecified.
- Investigator-initiated project—A project is investigator-initiated when an investigator develops the research program independently of the funding source.
- Manuscript—A peer-reviewed report that describes new research methods or findings or that summarizes and offers new insights about methods or findings previously reported in the literature.
- Proposal—A request for funding. Proposals may be for research, training, or infrastructure support. Infrastructure may include service centers, consulting cores, and activities generated by a data coordinating center. Proposals may or may not be funded, and they may require revisions before being funded.

Metrics

Proposals, grants, and contracts

- Numbers submitted, resubmitted, and funded
- Amounts requested, amounts awarded, and amounts allocated to BERD activities
- Funding sources and types (intramural or extramural; foundation, industry, or government agency; funding for infrastructure, research, or career development)
- BERD contributions to activities

Abstracts and presentations

- Number
- Inclusion of BERD practitioner as an author
- Forum for presentation (local, regional, national, or international)

Manuscripts

- Number
- Journal quality and impact metrics
- Inclusion of BERD practitioner as an author
- BERD contribution

Professional and Institutional Service

- Number of manuscripts peer reviewed
- Number of proposals reviewed for funding decisions
- Number of protocols reviewed for the protection of human subjects
- Number of protocols reviewed for scientific or methodological acceptability
- Participation on journal editorial boards

Definitions

- Participation on publications committees
 - Participation on data safety and monitoring boards or similar panels
 - Participation on institutional review boards or ethics committees
 - Participation in planning conferences
 - Leadership in professional societies
-

* BERD indicates biostatistics, epidemiology, and research design.

Table III

Definitions and metrics for the discovery of novel BERD-related methodologies *

Definitions[†]

- BERD-initiated discovery—The discovery of new methodology is BERD-initiated when the BERD practitioner identifies the need for new methods, as opposed to when an investigator presents with an identified need.
- Methodology—The approach used to collect, analyze, and interpret scientific data. Methodology is distinct from the scientific question itself. It may include the use of statistical approaches and measurement tools or involve the implementation and oversight of research.
- New methodology—A new method or the novel application of an existing method.
- Software—Software encompasses a wide range of electronic applications, including Web-based calculators and statistical tools, free-standing statistical tools, plug-ins and macros for commercial statistical software, and tools for randomization, group allocation, and sampling.

Metrics

New methodologies

- Initiator of the methodology (investigator or BERD practitioner)
- Type of methodology (statistical method, measurement method, etc.)

Proposals, grants, and contracts

- Numbers submitted, resubmitted, and funded
- Amounts requested and awarded
- Funding sources and types (intramural or extramural; foundation, industry, or government agency; funding for infrastructure, research, or career development)

Abstracts and presentations

- Number
- Forum for presentation (local, regional, national, or international)

Manuscripts

- Number
- Journal quality and impact metrics

Software

- Number
- Type of application developed (stand-alone program, plug-in, Web application, etc.)
- Use of application (number of licenses, downloads, citations, etc.)

Other means for the dissemination of information

- Educational activities (lectures, workshops, tutorials, etc.)[‡]
- Web sites
- Instruction manuals
- Reference books

* BERD indicates biostatistics, epidemiology, and research design.

[†] For additional definitions, see Tables I and II.

[‡] For metrics about educational activities, see Table I.