Quinlan, K. M., Kane, M., & Trochim, W. M. K. (2008). Evaluation of large research initiatives: Outcomes, challenges, and methodological considerations. In C. L. S. Coryn & M. Scriven (Eds.), *Reforming the evaluation of research. New Directions for Evaluation, 118,* 61–72.

5

Evaluation of Large Research Initiatives: Outcomes, Challenges, and Methodological Considerations

Kathleen M. Quinlan, Mary Kane, William M. K. Trochim

Abstract

The authors synthesize relevant literature and findings of evaluations of four large-scale, federally funded scientific research programs in the United States to identify desired outcomes of these types of programs, major evaluation challenges, and methodological principles and approaches. Evaluators face numerous contextual, political, and methodological challenges in evaluating big science. The authors propose that these may be addressed through participatory planning, such as concept mapping, triangulation of evidence, use of promising methodologies, and a systems approach. © Wiley Periodicals, Inc.

Biomedical research has undergone a significant transformation since the mid-1960s. In the United States, National Institute for Health (NIH) funding has been increasing since 1970 and doubled between 1998 and 2003 (Loscalzo, 2006). Concurrently, the nature, organization, and management of the scientific enterprise have changed (Edgerton, 1999; Nash & Stillman, 2003). Big science is now a significant portion of the NIH budget (more than $5 million) and is growing in popularity with Congress and the public as a way to call attention to specific health issues.

There are two major types of large-scale scientific research enterprise that stand in contrast to the traditional model of individual investigator-initiated awards. The first, centers of excellence programs, takes a multidisciplinary team approach focused on interaction between basic and clinical researchers to foster translational research. Major goals of center programs include offering more effective support for independently funded investigators, gaining increased attention to a program's research on the part of the center's home institution, recruiting established researchers to the program's area of interest, developing new investigators, expanding education of health professionals and the general public, and demonstrating state-of-the-art prevention, diagnosis, and treatment techniques (Manning, McGeary, Estabrook, & Committee for Assessment of NIH Centers of Excellence Programs, 2004). Research centers have also become popular with the public, advocacy organizations, and Congress, creating political pressure to establish new center programs to address specific diseases. Political interests, though, must be tempered by an understanding of when and how to make the best use of center grants (Manning et al., 2004).

Another type of large research initiative funded by the U.S. government is clinical research networks. Clinical research networks share some of the goals of center programs. However, they are focused on conducting clinical trials, which require coordination between multiple clinical centers, rather than direct interaction between basic and clinical investigators for translational research.

As government expenditures in scientific research have risen, pressure to demonstrate results from these investments also increases (Brainard, 2002a, 2002b; U.S. General Accounting Office, 2000). All government agencies are increasingly called on to demonstrate accountability, exemplified by the Government Performance and Results Act (GPRA) passed in 1993 (U.S. Office of Management and Budget, 1993). This effort led to a standardized governmentwide process for evaluating all federal programs, called the Program Assessment Rating Tool (PART), which, in turn, spawned a Web site for reporting results to the public (www.expectmore.gov). The National Institutes of Health Reform Act of 2006 (U.S. Congress, 2006) emphasized accountability for centers for excellence and included a specific requirement to report biennially on the performance and research outcomes of each center of excellence.

The confluence of these trends necessitates development of ways of evaluating the effectiveness of large-scale publicly funded scientific research enterprises. These large initiatives pose new management and evaluation challenges. The emergence of large initiatives requires assessment of a broader range of outcomes, including the social impact of the research (Smith, 2001). A recent Institute of Medicine report (Nash & Stillman, 2003) emphasizes measuring the technical and scientific output (such as data and research tools), benefits to the field, and the project's management and organizational structure, including staff performance, training, and retention. A 2004 IOM report recommends formal, external, retrospective review on a regular basis—at least every 5 to 7 years (Manning et al., 2004).

This chapter draws on the literature and the authors' experience with four projects to address three major questions on this topic: (1) What are the desired outcomes in large-scale, federally funded U.S. research initiatives?, (2) What are the major challenges in conducting these evaluations?, and (3) What methodologies are suggested by previous work?

Methodology

Using the lens of the three research questions just described, we examine four projects from the authors' portfolio of recent and current work, together with existing literature and documents. In all four projects, we developed an evaluation framework for a major research center program or clinical research networks program. Three of the four projects focus on the evaluation-planning phase of work, consistent with a finding from a 2004 study of 12 federal program evaluations that "the most critical factor in successful implementation of the evaluation design was the thoroughness of the design process" (Howell & Yemane, 2006, p. 234). The four projects are:

1. Centers for Disease Control and Prevention's Prevention Research Centers, or PRCs (Andersen et al., 2006)
2. National Cancer Institute's Transdisciplinary Tobacco Use Research Centers, or TTURCs (Stokols et al., 2003)
3. The National Institute for Allergy and Infectious Disease's Regional Centers of Excellence for Biodefense and Emerging Infectious Diseases Research Program (Concept Systems, 2007)
4. The National Institute for Allergy and Infectious Disease's Division of AIDS clinical research networks (National Institute of Allergy and Infectious Diseases [NIAID], 2006)

Each project used an integrated inquiry approach unified by concept mapping—a mixed-methods, structured group concept mapping methodology. Concept mapping (Trochim, 1989; Kane & Trochim, 2007) is a well-established social research method that combines familiar qualitative processes, such as brainstorming and sorting and rating of ideas with rigorous multivariate statistical analyses to create a shared conceptual framework. Each project involved a diverse range of stakeholders who defined success for their initiative, identified key evaluation questions, and weighed in on approaches that fit their context. The concept-mapping methodology is especially interesting in the context of evaluating large research efforts because it presents a rigorous structured approach that can be used effectively by scientists in the process of articulating the conceptual or logic model that underlies their endeavor—a major challenge in this type of evaluation.

Findings

The findings are organized by the three main research questions. Findings for each question integrate lessons learned from review of the four key projects and related literature.

Desired Outcomes. "Big science" initiatives are intended to achieve complex goals. Improved health outcomes are the ultimate intended impact of the public health research programs studied here. Given the long-term nature of that goal, we must look at the steps in the process that will lead from research to improved health outcomes. Scientific knowledge gained from research must be translated and applied in practice, policy, and service delivery to effect health outcomes. Indeed, a hallmark of center grant initiatives is the focus on translational science. Translating research may involve developing interventions, including prevention or treatment policies and practices or products such as vaccines or therapeutics. It may involve changes in public policy—an important outcome, particularly in the case of tobacco control. Finally, communication, public recognition, and uptake of new prevention and treatment approaches are essential elements in integrating research with practice. Advocacy organizations that lobby for creation of center programs expect the programs will lead to increased public attention to the topic.

Effective translation of scientific knowledge, though, depends on high-quality, strategically focused science. High-quality scientific research is traditionally evidenced by high-quality publications. In these projects, high-quality science involves addressing the most pressing questions with appropriate methods and designs that are feasible and logically connected to the rest of the research agenda. Stakeholders seek innovation in methods, models, technologies, and techniques. Finally, some expect flexibility in the management of the research portfolio, to enable responsiveness to emerging discoveries and scientific needs.

Joint ownership of, leadership for, and coordination of efforts in service of a broad strategic research agenda are particular concerns in these enterprises. Good science has always taken place within a community that shares and builds on prior knowledge, but big science brings the need for coordination of efforts into sharper focus.

Collaboration and coordination emerge as major elements of success across all the programs we studied. For many center grant initiatives, transdisciplinary integration is also a key factor for overall research success, adding a layer of complexity to collaboration.

Collaboration extends to broader communities of interest, which are defined according to the center or network. For instance, the mission of prevention research networks includes engaging community members as partners in applied research. In clinical trials research, respectful relationships with patient communities are critical to successful participant enrollment and retention. In other cases, interaction with local public health agencies is required to ensure community contribution. Community engagement is an important step throughout the research process. Involvement in the early phases of research helps ensure that the most pressing and appropriate questions are included in the research agendas, and it lays the foundation for successful translation and use at the end of a project.

Across the programs we investigated, recruitment of top-notch scientists and training of new investigators in the field are important goals. These efforts are described both as vital to immediate research productivity and important to creating an expanded cadre of experienced researchers in the field to support future advancements in the field. Success also depends on adequate financial resources, which participants in our projects described in two main ways. First, centers and networks seek to leverage resources from other sources, including host institutions, other government sources (noncenter or non-network funds), and industry. Second, typical grants management concerns about time lags, communication between program and accounting staff, and transparency of procedures are magnified with multi-institutional centers and networks. Not only must funding flow smoothly between the NIH and the core institution, but the centers or networks must also establish internal funds management practices that are accountable, transparent, and efficient across institutions.

Some of these large-scale research initiatives are specifically intended to build capacity and infrastructure. For instance, a goal of the Regional Centers of Excellence for Biodefense and Emerging Infectious Diseases Research is to build a web of regional laboratories to support emergency response in the event of an infectious disease emergency. Similarly, stakeholders expect the Prevention Research Centers to be an expert resource that will furnish technical assistance to public health organizations.

Efficient management is vital to success when dealing with the large sums of money, many people, substantial core facilities, and multi-institutional coordination associated with these enterprises. Scientific management has already been mentioned, but operations management is also required. With many centers engaged in similar activities, harmonizing operations and procedures and sharing key resources, including data, are also desirable. Through resource sharing and economies of scale, efficiency, and cost savings are often expected, particularly in the case of clinical research networks (Inventory and Evaluation of Clinical Research Networks, 2006b). Knowledge management is a subtext underlying overt collaboration and management themes.

Finally, our projects highlighted the need for effective management and leadership on the part of the granting agencies, as well as strong relationships between the grantees and the granting agency. Responsibility for success rests with both the grantees and the granting agencies.

Across various center and network research enterprises, there are common desired outcomes. There are also similarities in the pathways to those goals. Nonetheless, each initiative has its own unique purposes, histories, and contexts that must be understood and taken into account in creating a responsive evaluation design.

Challenges. These large biomedical research initiatives pose significant evaluation challenges. First, the long-term nature of scientific research makes it difficult to evaluate because there are various paths from basic research to human benefits (Manning et al., 2004; National Research Council, 1999; National Academy of Sciences, National Academy of Engineering, & Institute of Medicine, 1996, 2001).

As described earlier, these types of grants are designed to expand the human resources available in a field, in part by influencing the career paths of young investigators (Ailes, Roessner, & Feller, 1997). Yet it is difficult and time-consuming to track people over the period of time required to determine impact on career path (Manning et al., 2004).

These programs typically consist of many centers or networks. In our sample, the number of separate centers or networks ranges from 6 to more than 25. Each center or network typically has its own specialty area and operates under local conditions. This variability in context makes it difficult to compare centers or networks within a program meaningfully (Manning et al., 2004). Similarly, if results are aggregated across centers and networks within a program, it is difficult to identify appropriate comparisons or benchmarks at the program level.

In a related vein, although good center and network practices and operations contribute to achievement of positive outcomes, the effectiveness of clinical research networks and centers is also affected by funding agency input (Inventory and Evaluation of Clinical Research Networks, 2006a). Funding agencies must create funding flows, patterns, and support structures that will create conditions for grantee success. This interdependency suggests that an evaluation should be at the level of the enterprise as a whole, including attention to both the grantees and the granting agency. With many organizational layers contributing to the overall success of the enterprise, it is difficult to define the focus, level of detail, and boundaries of an evaluation.

A key question is whether center grant initiatives are a more effective approach to funding scientific research than other grant mechanisms. This question, though important to funding policy, is particularly difficult to answer. One approach is to focus on the unique mission of center grants. Another approach, which is fraught with the assumption of comparability of programs, is to focus on the common denominator of scientific quality and productivity. This second approach is further complicated by the difficulty of separating the effect of a center grant mechanism from other factors, such as highly talented and motivated individual investigators. Furthermore, when leveraging of funds is an explicit goal of a program, it is expected that center research will be supported not only by center funding but by other mechanisms and types of support (Manning et al., 2004; National Academy of Sciences et al., 1996, 2001). Something as simple as attributing a research paper to a particular grant can be difficult when multiple sources of support contributed to its completion. Paradoxically, it is just this type of synergy across researchers; institutions; and physical, financial, and intellectual resources that is a desirable outcome of this type of funding approach and may be at the heart of cost savings and efficiency.

This example illustrates the challenge involved in clarifying the stories we would like to tell about these initiatives. Programs must often balance tensions between competing goals, and this situation is no less so. Being aware of the connections, interdependencies, and tensions among goals is one step in the process of articulating program expectations and designing evaluations that will address them.

Peer review is a touchstone of scientific evaluation and should be incorporated into evaluation of large-scale research programs (Manning et al., 2004; National Academy of Sciences et al., 1996, 2001; Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine, 2000; Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, & Committee on Science, Engineering, and Public Policy, 2004). However, the size and scope of these research initiatives makes it difficult to find expert peer reviewers who are not involved with the program and do not have a conflict of interest (Manning et al., 2004; National Academy of Sciences et al., 1996, 2001).

Although the challenges described thus far are largely methodological, there are also significant barriers posed by tradition, context, and politics. Traditionally, investigators funded extramurally by the NIH work in academic institutions that have reward structures and evaluation criteria inconsistent with this new approach to scientific funding. A major goal of the big science approach is to cultivate new investigators, but large *collaborative* projects with long-term outcomes are not well rewarded in the individually oriented academic reward structure (Nash & Stillman, 2003). Transdisciplinary interaction, community engagement, and a focus on translational science are also goals that are not well supported by institutions in which researchers work.

Furthermore, substantive evaluation of many facets of the scientific enterprise is new for scientists and stands in contrast to the tradition of academic freedom and intellectual autonomy. Until now, granting agency expectations for accountability have been primarily administrative rather than substantial.

With such large-scale funding, decisions become highly politicized. Funded centers and networks put down roots in a place by virtue of capital investment in infrastructure, a large number of staff, and relationships with local community entities. When core facilities constitute a significant economic force on which a region or an institution depends, questions of its survival become politicized.

Power dynamics are also significant. Principal investigators of these large centers and networks control millions of dollars of funding, have considerable influence over the careers of many colleagues, and enjoy privileged positions in their home institution. They can become such a powerful force in their area of research that the typical power and authority relationship between grantor and grantee may be affected. Appreciating this political backdrop is useful in considering the incentives and disincentives of various stakeholders to participate in an evaluation.

Methodological Considerations. We do not attempt to catalog potential methods or measures here. Instead, we highlight several key themes in the development of this new branch of evaluation practice and illustrate promising approaches on the basis of our experience.

Participation. There are many stakeholders in the evaluation of large-scale scientific research, including federal agency staff, funded investigators and their staff, community members, other scientists, advocacy organizations, industry, and the public at large. To fully understand the program and its desired outcomes and construct meaningful measures that will be feasible to administer, these perspectives must be taken into account. Involving these stakeholders at all stages of the evaluation process helps ensure that the approach has the greatest utility. The example projects all began with a concept mapping process (Trochim, 1989; Kane & Trochim, 2007) that enabled thorough participation in creation of a shared conceptual framework of the expectations of the program. The result in each case was a collaboratively authored concept map that visually depicted participants' views of the success characteristics or outcomes of the program and the relationships among them. The map became the basis for a logic model of the program.

Attention to Consequences. Introducing an evaluation measurement into a system causes changes in practice in the system. In a complex and relatively new system, consequences may be difficult to predict. Therefore, we recommend engaging stakeholders in considering the potential consequences of a measure. Ideally, the evaluation process will help players in the system create conditions for success. Engaging stakeholders in this type of thinking helps to create a positive culture of evaluation.

Promising Methodologies. Across all of these large-scale research initiatives, stakeholders expect high-quality research. In the past few decades, bibliometric analysis has emerged as an important way of illuminating scientific influence and impact. Bibliometrics involves quantitative assessment of scientific publications, the works they cite, and citations of them. Several index variables enable comparison to baseline citation rate in the relevant literature. It builds on the system of peer review established within the scientific community (Osareh, 1996). Bibliometrics were used in the case of the Transdisciplinary Tobacco Utilization Research Centers project (Stokols et al., 2003).

Another emerging approach is use of survey techniques of staff and researchers. In the TTURCs project, the items generated by stakeholders as part of the concept mapping process were the starting point for developing and validating a comprehensive survey that covered the whole range of identified outcome areas (Stokols et al., 2003).

Other researchers (Corley, Melkers, & Johns, 2006; Vonortas & Malerba, 2005; Zuckerman & Kupfer, 2005) have used social network analysis as a tool in evaluating collaboration in these types of research program.

Historically, science has been evaluated by assessing the scientific quality of the work, largely through peer review of research proposals and publications (Jefferson & Godlee, 1999; Kostoff, 1994a, 1994b, 1995). Virtually all recent studies of this topic recommend peer review, with notable enhancements and modifications that take into account the current context of science (Manning et al., 2004; National Academy of Sciences et al., 1996, 2001; Committee on Science, Engineering, and Public Policy et al., 2000; Committee on Facilitating Interdisciplinary Research et al., 2004). The judgments of external, impartial experts with the knowledge and experience to understand the research must be balanced with other sources of data. To evaluate research in progress, the Committee on Science, Engineering, and Public Policy et al. (2000) recommend that experts (both scholars and users) review quantitative data on the basis of the criteria of quality, relevance, and leadership. They further recommend use of a process of international benchmarking (Committee on Science, Engineering, and Public Policy et al., 2000), using a panel of non-U.S. and U.S. experts to assess "the relative position of US research today, the expected relative position of US research in the future, and the key factors influencing relative US performance" (National Academy of Sciences et al., 2001, p. 11). The interdisciplinary focus of many large research initiatives requires inclusion of researchers with interdisciplinary expertise, as well as researchers with expertise in the relevant disciplines (Committee on Facilitating Interdisciplinary Research et al., 2004).

Assembling a Body of Evidence. No single analysis is likely to establish definitively the effects of such complex programs. Therefore designs should incorporate a variety of measures that, taken together, can demonstrate a pattern of results that will enable reasonable causal inference. When more than one type of evidence is gathered for the same construct, as with reports from investigators and site-visit-report and center-profile data, evaluations should cross-check these data sources (Committee on Science, Engineering, and Public Policy et al., 2000). The evaluator, like an attorney, assembles the case so that each piece of evidence is woven together with other pieces of evidence to create a story. Expert panels, as we have described, may be the jury who weighs in on the case, adding their interpretations. The conceptual models of the program, such as the concept-map framework and resulting logic model described earlier, offer a vehicle for integrating results from multiple inquiries.

Systems Approach. Typically, evaluation is done as a single ad hoc event. In contrast, our experience suggests the need for a systems approach to evaluation. A systems approach seeks to build in rigorous evaluation throughout the funding process and life cycle, to ensure a seamless integration with other grant activities and give feedback at key points to inform decision making by various users. There is little precedent for systems approaches to evaluation of large scientific research endeavors (see as exceptions Stokols et al., 2003; and National Academy of Sciences et al., 1996). However, the size, scope, and goals of these large-scale scientific research initiatives suggest this need. Building feedback mechanisms into the life of the grant is philosophically and practically consistent with the funding mechanisms. Doing so also presents the challenge of evaluating the program as a whole, which includes both the funder and the grantees.

Conclusion

Taken together, these four projects enrich our understanding of the U.S. context of the evaluation of large-scale research initiatives. They highlight key questions and outcome domains that specific center programs and clinical research networks must address. They also point to challenges that must be addressed by evaluators.

Evaluators must adjust their role to fit the unique context of scientific research, drawing on their evaluation methodology and research expertise and a rich understanding of the unique evaluation context of scientific research. They must offer expertise in designing an evaluation that is sensitive to the scientific context. They also have to be able to offer assistance and advice in integrating evaluation into existing systems, coordinate and aggregate expert judgments, compare expert judgment to other sources of data that address the same evaluation questions, and help to ensure that the results become part of the feedback to pertinent users in the system.

References

Ailes, C., Roessner, D., & Feller, I. (1997). *The impact on industry of interaction with engineering research centers* (Final report prepared for the National Science Foundation). Arlington, VA: SRI International. Retrieved August 8, 2006, from http://www.sri.com/policy/csted/reports/sandt/erc/

Andersen, L., Gwaltney, M., Sundra, D., Brownson, R., Kane, M., Cross, A., et al. (2006). Using concept mapping to develop a logic model for the prevention research centers program. *Preventing Chronic Disease, 3*(1). Retrieved April 3, 2007, from www.cdc.gov/pcd/issues/2006/jan/05_0153.htm

Brainard, J. (2002a, June 21). New director discusses NIH's massive budget. *Chronicle of Higher Education.* Retrieved June 13, 2007, from http://chronicle.com/weekly/v48/i41/41a02503.htm

Brainard, J. (2002b, March 29). New science measures related by OMB. *Chronicle of Higher Education*. Retrieved June 13, 2007, from http://chronicle.com/weekly/v48/i29/29a02502.htm

Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, Institute of Medicine, & Committee on Science, Engineering, and Public Policy. (2004). *Facilitating interdisciplinary research*. Washington, DC: National Academies Press.

Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2000). *Experiments in international benchmarking of US research fields*. Washington, DC: National Academies Press.

Concept Systems. (2007). *Development of an evaluation framework for DMID's regional centers of excellence for biodefense and emerging infectious diseases program: Concept mapping summary report*. Ithaca, NY: Concept Systems.

Corley, E., Melkers, J., & Johns, K. (2006, May). *Layered and evolving networks: Innovative evaluation methods for interdisciplinary research in university-based research centers*. Paper presented at Atlanta Conference on S&T Policy, Atlanta.

Edgerton, D. E. H. (1999). Before big science: The pursuit of modern chemistry and physics, 1800–1940. *Annals of Science, 56*(1), 100–107.

Howell, E. M., & Yemane, A. (2006). An assessment of evaluation designs: Case studies of twelve large federal evaluations. *American Journal of Evaluation, 27*(2), 219–236.

Inventory and Evaluation of Clinical Research Networks. (2006a). *Best practices study final report*. Rockville, MD: WESTAT.

Inventory and Evaluation of Clinical Research Networks. (2006b). *Core and descriptive survey final report*. Rockville, MD: WESTAT.

Jefferson, T., & Godlee, F. (1999). *Peer review in health sciences*. London: British Medical Journal Publishing Group.

Kane, M., & Trochim, W. M. K. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks, CA: Sage.

Kostoff, R. N. (1994a). Assessing research impact—Federal peer-review practices. *Research Evaluation, 18*(1) 31–40.

Kostoff, R. N. (1994b). Quantitative qualitative federal research impact evaluation practice. *Technological Forecasting and Social Change, 45*(2), 189–205.

Kostoff, R. N. (1995). Research requirements for research impact assessment. *Research Policy, 24*(6), 869–882.

Loscalzo, J. (2006). NIH-sponsored clinical trials worth the cost, Reuters Health Information: The NIH budget and the future of biomedical research. *New England Journal of Medicine, 354*(16), 1665–1667.

Manning, F. G., McGeary, M., Estabrook, R., & Committee for Assessment of NIH Centers of Excellence Programs. (2004). *National Institute of Health extramural center programs: Criteria for initiation and evaluation*. Washington, DC: Institute of Medicine, Board on Health Sciences Policy.

Nash, S. J., & Stillman, B. W. (2003). *Large-scale biomedical science: Exploring strategies for future research*. Washington, DC: National Academies Press.

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (1996). *An assessment of the National Science Foundation's science and technology centers program*. Washington, DC: National Academies Press.

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2001). *Implementing the Government Performance and Results Act for research: A status report*. Washington, DC: National Academies Press.

National Institute of Allergy and Infectious Diseases (NIAID). (2006). *NIAID announces leadership for newly restructured HIV/AIDS clinical trials networks, 2006*. Retrieved August 17, 2006, from http://www3.niaid.nih.gov/news/newsreleases/2006/leadership.htm

National Research Council. (1999). *Evaluating federal research programs: Research and the Government Performance and Results Act*. Washington, DC: National Academies Press.

Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature. *Libri, 46*(3), 149–158.

Smith, R. (2001). Measuring the social impact of research: Difficult but necessary. *British Medical Journal, 323,* 528.

Stokols, D., Fuqua, J., Gress, J., Harvey, R., Phillips, K., Baezconde-Garbanati, L., et al. (2003). Evaluating transdisciplinary science. *Nicotine and Tobacco Research, 5,* S-1, S21–S39.

Trochim, W. (1989). An introduction to concept mapping for planning and evaluation. In W. Trochim (Ed.), *Special Issue of Evaluation and Program Planning, 12,* 1–16.

U.S. Congress. (2006). *National Institutes of Health Reform Act of 2006. 109th Cong., H.R. 6164*. Retrieved March 9, 2007, from http://www.govtrack.us/congress/billtext.xpd?bill=h109–6164

U.S. General Accounting Office. (2000). *NIH research: Improvements needed in monitoring external grants* (GAO-HEHS-AIMD-00–139). Washington, DC: U.S. General Accounting Office.

U.S. Office of Management and Budget. (1993). *Government Performance Results Act of 1993. Executive Office of the President of the United States, 1993.* Retrieved August 17, 2006, from http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m.html

Vonortas, N. S., & Malerba, F. (2005, October). *Using social network methodology to evaluate research and development programs.* Paper presented at the joint conference of the Canadian Evaluation Society and the American Evaluation Association, Toronto, Ontario.

Zuckerman, B. L., & Kupfer, L. (2005, October). *Social network-based design of collaborative research program evaluation.* Paper presented at the joint conference of the Canadian Evaluation Society and the American Evaluation Association, Toronto, Ontario.

KATHLEEN M. QUINLAN is director of research and a senior consultant at Concept Systems.

MARY KANE is president and chief executive officer of Concept Systems.

WILLIAM M. K. TROCHIM is professor of policy analysis and management at Cornell University, where he is also director of evaluation for extension and outreach.