

Evaluation Guidelines for the Clinical and Translational Science Awards (CTSAs)

William M. Trochim, Ph.D.¹, Doris M. Rubio, Ph.D.², Veronica G. Thomas, Ph.D.³, and the Evaluation Key Function Committee of the CTSA Consortium

Abstract

The National Center for Advancing Translational Sciences (NCATS), a part of the National Institutes of Health, currently funds the Clinical and Translational Science Awards (CTSAs), a national consortium of 61 medical research institutions in 30 states and the District of Columbia. The program seeks to transform the way biomedical research is conducted, speed the translation of laboratory discoveries into treatments for patients, engage communities in clinical research efforts, and train a new generation of clinical and translational researchers. An endeavor as ambitious and complex as the CTSA program requires high-quality evaluations in order to show that the program is well implemented, efficiently managed, and demonstrably effective. In this paper, the Evaluation Key Function Committee of the CTSA Consortium presents an overall framework for evaluating the CTSA program and offers policies to guide the evaluation work. The guidelines set forth are designed to serve as a tool for education within the CTSA community by illuminating key issues and practices that should be considered during evaluation planning, implementation, and utilization. Additionally, these guidelines can provide a basis for ongoing discussions about how the principles articulated in this paper can most effectively be translated into operational reality. *Clin Trans Sci* 2013; Volume #: 1–7

Keywords: evaluation, translational research, CTSA

Introduction

Translational research is a critically important endeavor in contemporary biomedical research and practice. Yet, there have long been considerable stumbling blocks to successful translational research efforts in science, medicine, and public health. Several years ago, a seminal paper published in *The Journal of the American Medical Association* stressed, “Without mechanisms and infrastructure to accomplish this translation in a systematic and coherent way, the sum of the data and information produced by the basic science enterprise will not result in tangible public benefit.”⁶ To address this major concern, through discussions with deans of academic health centers, recommendations from the Institute of Medicine, and meetings with the research community, the National Institutes of Health (NIH) recognized that a broad reengineering effort was needed to create greater opportunity to catalyze the development of a new discipline of clinical and translational science. This resulted in NIH launching the Clinical and Translational Science Awards (CTSA) in October 2006. The program, supported initially by the National Center for Research Resources (NCRR) and then by the National Center for Advancing Translational Sciences (NCATS), currently funds a national consortium of 61 medical research institutions in 30 states and the District of Columbia that are seeking to transform the way biomedical research is conducted, speed the translation of laboratory discoveries into treatments for patients, engage communities in clinical research efforts, and train a new generation of clinical and translational researchers. The CTSA program, at a cost of approximately a half billion dollars per year, is part of a larger 21st century movement to develop a discipline of clinical and translational science.

An endeavor as ambitious and complex as the CTSA program requires high-quality evaluation. The program needs to show that it is well implemented, efficiently managed, and demonstrably effective. Evaluation is key to achieving these goals. Without

appropriate evaluative data and assessment systems, it would be difficult to guide development of policies for CTSAs, in general, and CTSA operations more specifically. Evaluation of the CTSA program can provide the prospective and retrospective information necessary to direct its course and to assess the degree to which the program is accomplishing its goals.

The purposes of this paper are to present an overall framework for evaluating the CTSA program and to offer policies to guide the evaluation work. The term *CTSA program* refers here to the entire initiative that encompasses 61 sites across the United States and a national consortium consisting of representatives of the sites and of the NIH. This document is not intended to be a series of prescriptive requirements. Rather, it is intended to provide general guidance on evaluation in the CTSA context, to discuss the critically important role of evaluation, and to present recommendations designed to enhance the quality of current and future CTSA evaluation efforts. While much of what is addressed in the document may be directly generalizable to translational research efforts outside of the CTSAs or to other types of large multicenter grant initiatives, the focus of the document is on the CTSA context.

The guidelines provided in this paper are intended to offer recommendations for the myriad ways that CTSA evaluations can be accomplished given the range and complexity of individual and collective CTSA evaluative efforts. These recommendations are based upon the best practices from the discipline and profession of evaluation. The leading professional association in the field, the American Evaluation Association (AEA), has through its Evaluation Policy Task Force (EPTF) produced the document “*An Evaluation Roadmap for a More Effective Federal Government*” (The “AEA Roadmap”) that we have learned from and modeled in developing CTSA evaluation guidelines. The authors, who are members of the CTSA Evaluation Key Function Committee’s National Evaluation Liaison Workgroup, received consultation

¹Department of Policy Analysis and Management, Cornell University, Ithaca, New York, USA; ²Data Center, Center for Research on Health Care, Division of General Internal Medicine, Department of Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; ³Department of Human Development and Psychoeducational Studies, Howard University, Washington, DC, USA.

Correspondence: William M. Trochim (wmt1@cornell.edu)

DOI: 10.1111/cts.12036

and feedback from the AEA's EPTF in the formulation of the guidelines provided in this paper. Additionally, considerable input was obtained from members of the Evaluation Key Function Committee of the CTSA Consortium.

The CTSA Context

The CTAs, awarded by the NIH, constitute one of the most ambitious and important endeavors in biomedical research and practice in the early part of the 21st century. The CTSA program was born out of the NIH Roadmap trans-NIH initiatives designed to accelerate the pace of discovery and improve the translation of research findings to improve healthcare. Launched in 2006 by the NIH, the program supports a national consortium of medical research institutions that seek to transform the way biomedical research is conducted. The goals of the program are to accelerate the translation of laboratory discoveries into treatments for patients, to engage communities in clinical research efforts, and to train a new generation of clinical and translational researchers. At its core, the CTSA program is designed to enable innovative research teams to speed discovery and advance science aimed at improving our nation's health, tackling complex medical and research challenges, and turning discoveries into practical solutions for patients.

Unlike most other large initiatives of the NIH, the CTSA initiative included evaluation efforts at its outset. The NIH required each CTSA institution to develop an evaluation program and to undertake site-level evaluations. It also required a national evaluation of the entire CTSA initiative to be conducted by an external evaluator.

The Importance of Evaluation

For purposes of this document, we use the definition of evaluation that was developed in 2008 by Patton: evaluation is the “systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase understanding.”¹

Evaluation activities fall on a continuum from program evaluation to evaluation research. On the program evaluation end of the continuum, examples of activities are program model development, needs assessment, tracking and performance monitoring, continuous quality improvement, and process and implementation analysis. On the evaluation research end of the continuum, examples are precise measurement, testing of program theories, and the assessment of outcomes and impact (including quasi-experimental and randomized experimental designs). Evaluation is itself an established field of inquiry and is essential for both organizational management and learning. It helps organizations anticipate needs, articulate program models, and improve programmatic decisions.

The NIH explicitly recognized the critical importance of evaluation for the CTAs by requiring in the Request For Applications (RFA) that it be integrated into all levels of the endeavor. Specifically, RFAs have required that each CTSA have an evaluation core that assesses administrative and scientific accomplishments, conducts self-evaluation activities, and participates in a national evaluation.

Recommendations

The following sections provide general observations and recommendations that can guide how CTSA evaluation is

framed within selected topical areas. The recommendations are not prescriptive; they are meant to offer broad guidance for CTSA evaluation. They are based on the observations and experience of evaluators currently engaged in the CTAs. Some of the recommendations may have direct implications for policy or action. Others have the intention of clarifying a conceptual concern or encouraging thinking about a complex evaluation challenge.

Scope of CTSA Evaluation

The range and complexity of evaluation questions and issues that need to be addressed in CTSA contexts are both exciting and daunting. The breadth of purpose and scope constitutes one of the major challenges for CTSA evaluation.

Evaluation should engage stakeholders in all phases of the evaluation

It is critical to identify and engage the stakeholders from the beginning of the evaluation. This will ensure that the input, guidance, and perspectives of stakeholders are incorporated in all phases of the evaluation, and it will also ensure that the stakeholders are kept informed and are able to utilize the findings. Stakeholders may be engaged through an advisory panel that has input into identifying evaluation goals, developing an evaluation strategy, interpreting findings, and implementing recommendations.

Evaluation should be an integral part of program planning and implementation

A common misunderstanding is that evaluation is simply an “add-on” to program activities, rather than an integral part of a program's structure. Evaluation serves a program best when it is coordinated with program development and is ongoing and responsive. Evaluative thinking is a critical element in program planning. Program design is significantly enhanced by clarifying goals and objectives, specifying logical frameworks for interventions, and considering key metrics in advance. Ensuring timely, high-quality evaluation is a broad-based responsibility of all key stakeholders, including policymakers, program planners and managers, program staff, and evaluators. It is especially important that CTSA program leaders take ownership of evaluation in an active way, including ensuring that evaluation addresses questions of central importance to the leadership and to other key stakeholders. In the complex environment of a CTSA site, evaluation should be an ongoing function distributed across all cores.

A balanced set of evaluation activities and methods needs to be encouraged both at the CTSA site level and at the national consortium level

No one type of evaluation will meet the needs of the multiple stakeholders involved in the CTAs. For principal investigators and administrators, the primary interest is likely to be in process and implementation evaluation that helps them manage their CTAs. For NIH staff, the priority is on standardized metrics and cross-cutting analyses that enable aggregation, provide evidence of scientific productivity, and offer a clear picture of how the CTSA initiative is performing in a way that can guide future program direction. For Congress and the public, the primary focus is on compelling evidence of the impact of the CTAs on the health of the public. No one evaluation approach will meet all needs.

The challenge will be to find a combination of approaches that meet the variety of information needs in an efficient and cost-effective manner.

CTSA evaluation should be prospective as well as retrospective and should be ongoing and connected to program management

Many people think of evaluation only as a retrospective endeavor that looks back on some program and assesses the degree to which it worked to achieve its goals. But evaluation is necessarily more than that. For evaluation to be effective, the evaluators need to develop a clear model of an intervention (such as a logic model) that shows the major activities, outputs, and short-, mid-, and long-term outcomes and illustrates how all of these are interconnected in an organizational context. This type of modeling activity is prospective and intimately tied to program planning and management. Effective evaluation is ubiquitous; it occurs before, during, and after a program has been delivered or implemented. It is more appropriately viewed as an organizational feedback and learning mechanism than as a separate activity that is disconnected from the everyday functioning of the program. CTSA evaluation should be ongoing and connected to program management.

CTSA evaluation should involve a combination of internal and external approaches

Internal evaluation typically is conducted by organizational staff, emphasizes the provision of feedback on the functioning of the program, and is used to improve the program's management. External evaluation is usually conducted by evaluators who are not directly involved in the administration of the program and is used primarily to assess the effects of the program. However, these two approaches are intimately connected. For example, local CTSA evaluators are well-positioned to assist in the collection of cross-CTSA standardized data that can be aggregated as part of an external national evaluation. For an effort such as this, clearly defined data collection protocols are essential to ensure objective and accurate data. Wherever possible, data collected for cross-institution aggregation should be made available for internal local evaluation purposes as well.

The highest professional standards for evaluation should be followed

The field of evaluation has several sets of standards to guide evaluators in their professional work. For instance, the AEA has the *Guiding Principles for Evaluators*,² a document that covers the topics of systematic inquiry, competence, integrity and honesty, respect for people, and responsibilities for general and public welfare. In addition, a diverse group of professional associations has developed *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*,³ which is directly relevant to CTSA evaluation. The CTSA evaluation endeavor should consciously follow the highest professional standards, with evaluation policies constructed to support these standards.

CTSA evaluation needs to combine traditional approaches with innovative and cutting-edge approaches

CTSAs pose significant and unique challenges that require novel and innovative evaluation methods and approaches. CTSA evaluation itself is a learning endeavor. It has the potential for significantly enhancing the scope of the discipline of evaluation, especially with respect to the evaluation of large, complex scientific

research initiatives, about which much still needs to be learned. CTSA evaluators can be worldwide leaders in this endeavor, but it will require an institutional and organizational commitment of energy and resources for the work to make the cutting-edge contribution that it is capable of making.

Structural and Organizational Issues

The CTSA program and its sites should establish and continuously improve a formal evaluation planning process

For any high-quality program, including the CTSA program, it is important to develop and improve a formal evaluation planning process, one that periodically (e.g., annually) lays out a multiyear plan for evaluation that is both strategic and mission-oriented. The CTSA sites and multiple programs within them are continually evolving through different stages of development. The evaluation approaches used in early developmental stages are inappropriate for the later, more mature phases in which summative approaches that focus on outcomes are used.

Because the CTSA is a 5-year renewable award, individual site-level evaluation subprojects may be active during different time periods. However, the overall evaluation plan should link clearly to a larger strategic vision and include a model that describes the current context and assumptions of the CTSA program. Each individual evaluation proposal should include the evaluation questions being addressed and the sampling, measurement, design, and analysis that will be conducted, and it should outline how the results will be disseminated and utilized. In addition, the plan should include potential evaluation ideas and approaches to be explored or piloted for the future. Because of the varied and complex nature of the CTSA, such a plan would likely include diverse evaluation approaches, ranging from feasibility and implementation studies to cost-benefit and return-on-investment studies to process and outcome evaluations. Approximate timelines and associated skills and resource needs should be specified and prioritized. Such a plan is important in any evaluation but is especially important with a complex effort like the CTSA program.

CTSA evaluation should address the entire range of translational research from basic discovery to effects on the health of the public

Translational research encompasses the research-practice continuum from basic discovery through clinical testing to translation into practice and, ultimately, to effects on the health of the public.⁴⁻⁸ The success of the entire endeavor depends on how well the system addresses its weakest links. While the CTSA may elect strategically to focus on some parts of translation more than others, they should ensure that their evaluation portfolio includes a balance of activities that can address the entire scope of the translational process. For example, detailed case histories of successful translational research efforts can document the key milestones and pathways taken across the translational continuum and can contribute considerably to our understanding of subparts of the larger process.

The CTSA Consortium should work collaboratively with the national CTSA evaluators to identify and pilot-test a small, rigorous set of standard definitions, metrics, and measurement approaches for adoption by all CTAS

The development of standard definitions and metrics has received considerable attention by members of the CTSA Evaluation

Key Function Committee's National Evaluation Liaison Group and other CTSA stakeholders. Discussions have highlighted that definitions are needed for key metrics that would enable stakeholders to determine which researchers, projects, or grants benefited directly from the CTSA and to assess collaboration and interdisciplinarity.

The development of standard metrics requires consensus on the definition and measurement of the metrics. With 61 CTASs throughout the country, identifying and establishing standard metrics is a daunting task requiring thoughtful deliberation. However, the potential benefits are considerable. The Biostatistics, Epidemiology, and Research Design (BERD) Key Function Committee formed a subcommittee on evaluation. Starting with the metrics proposed in each successful CTSA grant application, this subcommittee identified and operationalized key metrics for BERD units within CTASs.⁹ This pioneering work is a model that other CTSA key function committees can emulate to identify key metrics in their own areas of focus.

A set of standard quantitative and qualitative metrics is crucial for cross-sectional and longitudinal comparisons across different CTASs and is needed to facilitate smoother interaction between the national and local evaluation groups. However, standard metrics are only one building block for evaluation. They do not by themselves constitute effective evaluation. There are some important outcome areas for which simple, quantitative metrics will not be possible or sufficient, and metrics must always be interpreted within the context in which they are produced.

CTSA evaluation should be an integrated multilevel systems endeavor

Evaluation of the CTSA program occurs for many purposes and on many levels, ranging from the smallest CTSA unit at the site level to the largest CTSA unit at the national level, and should include site-wide and program-wide evaluations of entities such as CTSA cores and the CTSA Consortium key function committees.

In complex programs like the CTASs, stakeholders should be mindful of the different levels of evaluation activity and be clear about roles, responsibilities, and strategies for coordination. They should understand that planning and communication are needed to ensure that evaluation activities are complementary and mutually supportive. Disparate stakeholders located at different levels in the organizational structure are likely to hold different interests in and expectations for evaluation. For example, in the organizational structure, the higher an entity moves (e.g., toward the national cross-CTSA level), the greater is the focus on external evaluation, longer-term outcomes, and policy-level issues. The lower an entity moves (e.g., toward the specific CTSA key function level), the greater is the emphasis on internal evaluation, process monitoring and improvement, and shorter-term outcomes and management issues.

The CTSA program should be proactive and strategic regarding how to coordinate and integrate evaluation conducted at different organizational levels

The CTASs involve both local and national evaluation components and need to think strategically about how to maximize their relationship. The situation is complicated by the fact that the needs of local and national evaluations are convergent in some areas and divergent in others. To share information on evaluation plans and to reduce confusion and response burden, there must be continuing dialogue between these levels. For example,

local evaluators must continue to be represented on a national evaluation advisory group. Additional opportunities should be sought for meaningful collaboration that contributes to the knowledge of clinical and translational research and of evaluation.

Evaluation Methodology

Tracking and monitoring should be considered necessary but not sufficient components of CTSA evaluation

Tracking and monitoring are integral parts of the CTSA evaluation portfolio, but they do not in and of themselves provide all that is needed to understand how well the CTASs are functioning, the extent to which they are making progress toward meeting program goals, and how they might be improved. Examples of data that are tracked by the CTASs are the number of clinical and translational science pilot studies that are funded, the number of CTSA investigator manuscripts that are published in peer-reviewed journals, and the number of individuals from underrepresented groups who enroll in clinical and translational science educational offerings. While tracking information is certainly valuable, a preoccupation with this type of data can result in overemphasis on measuring what is easy and accessible, rather than focusing on variables that have a wider impact on the transformation of clinical and translational science within, across, and beyond the CTSA Consortium.

CTSA evaluation should include both process evaluation and outcome evaluation

Process evaluation is distinct from outcome evaluation. Process evaluation begins in the early stages of program development and continues throughout the life of the program. Its primary goal is to provide information that will guide program improvement. In contrast, outcome evaluation is undertaken to provide a summary judgment about the extent to which the program has reached or is making progress toward reaching its stated goals. A premature emphasis on outcome evaluation for a newly developing program is no more sensible than the use of process evaluation approaches long after processes have been examined, revised, and standardized. While both types of evaluation are important to a comprehensive assessment of the CTSA program, each type is emphasized at different points in a program's life cycle. Taken together, they can help researchers, program managers, and policy-makers clarify goals, determine whether goals are attained, and understand the factors that facilitate or inhibit goal attainment.

CTSA evaluation should include a balanced portfolio of methods that encompass local CTSA variations and nationally standardized approaches

Local CTASs are continually trying new evaluation approaches or adapting existing ones to address locally relevant questions or unique contextual circumstances. This experimental approach to evaluation is appropriate and can act as an incubator and testing ground for potentially useful and generalizable evaluation plans. Cross-CTSA standardization can enable aggregation and analyses across the CTASs. Both local variation and national standardization are needed. While some stakeholders might argue that all CTSA evaluations should be nationally standardized, that would not allow for the considerable variation of purposes and approaches of the different CTSA sites. The CTSA Consortium should seek an appropriate balance between local variation and national standardization of evaluation.

The CTSA evaluation portfolio should incorporate a mix of qualitative and quantitative methods

No single method can assess the ultimate value of the CTSA program or decipher its complexities. Evaluators use a combination of qualitative and quantitative methods to achieve the advantages and minimize the potential disadvantages of each method. The mixed-methods approach is well-suited for capturing the full complexity of the CTSA local, regional, and national activities and providing evidence to determine whether the CTSA program is achieving its intended goals. By building on the strengths of each type of data collected and minimizing the weakness of any single evaluation approach, mixed methods can increase both the validity and reliability of data and results.

The CTSA evaluation portfolio should involve piloting and experimenting with new evaluation methods or variations on methods

While there are a number of traditional evaluation approaches that can be applied to the CTAs, cutting-edge evaluation approaches are necessary to assess some of the more innovative aspects of the program. These approaches would include exploring developmental evaluation¹⁰ and systems evaluation,^{11,12} as well as conducting research on the evaluation process itself.

Utilization of Evaluation

Evaluation plans need to address how evaluation results will be used

There are multiple ways in which learning from evaluations can benefit decision makers. These include providing a deeper understanding of a policy or management problem, recommending strategies to modify or improve a program or policy, providing information on program performance and milestones achieved, illuminating unintended consequences of a policy or program, and informing deliberations regarding the allocation of resources. Currently, there is an urgent push for decision makers, including members of Congress, to become informed consumers of evaluation data when they weigh considerations and make decisions. The AEA's 2010 document called *An Evaluation Roadmap for a More Effective Government*¹³ stresses the need to make evaluation integral to managing government programs at all stages, from planning and initial development through startup, ongoing implementation, and appropriations and, ultimately, to reauthorization. Evaluation results are useful for decision makers because they can help inform practice. For example, the results can be used to refine programs, enhance services, and, in some cases, eliminate program activities that are not effective. Moreover, the evaluation results can prove useful to individuals who conduct evaluation research because these individuals can learn from the results and use them to develop better and more innovative methods and approaches to evaluation.

The CTSA program should assess the degree to which evaluations are well-conducted and useful in enhancing the CTSA endeavor

The field of evaluation uses the term *meta-evaluation* to refer to efforts that assess the quality and effects of evaluation with the aim of improving evaluation efforts. Meta-evaluation would be helpful for providing feedback to the CTSA Consortium about how its evaluation efforts are proceeding. It would also be helpful for assessing the degree to which stakeholder groups use the

evaluation and perceive it as beneficial. Meta-evaluation should be incorporated into formal CTSA evaluation planning.

CTSA evaluation needs to be open, public, and accessible

In today's policy debates, there is much discussion about transparency and the use of technology to make the government more accessible and visible to the public. Transparency in the federal sector serves multiple aims, including promoting greater accountability, building public trust and confidence, and creating a more informed citizenry. In the case of the CTSA program, there is a perceived delicate balance involved in maintaining transparency and public records of evaluation while allowing for competition and entrepreneurship among CTAs. In addition to ensuring that stakeholders have access to the CTSA information they need, transparency generates accountability at both local and national levels and adds an important level of credibility to the entire evaluative enterprise.

Evaluation Policy

An evaluation policy is any rule or principle that a group or organization uses to guide its decisions and actions with regard to evaluation.¹⁴ All entities that engage in evaluation, including government agencies, private businesses, and nonprofit organizations, have evaluation policies, and many have adopted quality standards to guide their evaluations. Evaluation policies can be implicit and consist of ad hoc principles or norms that have simply evolved over time. Alternatively, they can be explicit and written. Written evaluation policies or guidance should address a number of important topics, such as evaluation goals, participation, capacity building, management, roles, processes, methods, use, dissemination of results, and meta-evaluation. Developing explicit written evaluation policies clarifies expectations throughout the system, provides transparency, and delineates roles.

General written evaluation policies should be developed for the CTSA program and its sites

The CTAs already have the beginnings of evaluation policies from the previous RFA requirements for evaluation. These requirements were groundbreaking for the NIH and represent one of the most ambitious evaluation undertakings ever attempted for a large, complex federal grant initiative. The RFA requirements are not formally structured as policies, and there are numerous implicit assumptions that warrant clarification both at the national consortium level and within local CTAs. For instance, it is not always clear who should be responsible for collecting specific types of data; how cross-center data collection efforts will be managed; how evaluation projects will be proposed, accepted, or rejected; and how evaluation results should be reported. The Evaluation Key Function Committee should work with the national consortium leadership to clarify current policies in writing. Evaluation policy is a high priority for many individuals engaged in evaluation leadership outside the CTSA community,¹⁴ and their expertise and products may help in developing appropriate policies for the CTAs.

Evaluation policies need to be developed collaboratively with the aim of providing general guidance rather than specific requirements

The CTAs are a collaborative network. At the national level, written evaluation policies should be developed to give local CTAs general guidance while allowing them the flexibility they

need to function effectively and encouraging them to provide greater evaluation policy specificity at the local level. For example, rather than developing national policies for how the evaluation function is to be organized at the local level, who should be responsible for local data collection, and so on, the national consortium should call on each CTSA to develop its own written policies to address these issues in a locally relevant manner.

Evaluation policy should encourage participatory and collaborative evaluation at all levels

Professional evaluating standards require that assessments be sensitive to the perspectives and values of multiple stakeholder groups. CTSA policy at the national level should embrace diverse perspectives within and across various levels of CTSA efforts.

Evaluation Capacity and System Development

As the NIH plans for a robust evaluation foundation, it is important to keep in mind that the field of evaluation, like that of other sciences, needs to be nurtured, supported, and challenged to ensure continued growth and capacity. It is especially important for individuals involved in the CTSA initiative to understand the range of evaluation approaches, designs, data collection, data analysis, and data presentation strategies that might be applied to deepen an understanding of the progress that the initiative is making. Building capacity in this area can and should take many different forms.

The NIH should encourage ongoing professional development and training in evaluation as appropriate at all organizational levels and bring to bear the proper mix of evaluation skills to accomplish evaluation

Evaluation professionals and evaluation managers should have the opportunity to pursue continuing education to keep them abreast of emerging designs, techniques, and tools. In addition, it is important to ensure that CTSA program leaders, CTSA program officers, principal investigators, and staff at individual CTSAs receive some grounding in evaluation so they can better understand how to maximize the evaluation process and findings to reach their goals. Recognizing that CTSA evaluation requires a broad array of skills and methods, each CTSA should identify a team that can effectively plan and carry out evaluations. This team should include a trained evaluation professional who can bring specialized evaluation skills to bear and, as needed, should include other professional staff with relevant skills.

The CTSA program should leverage its network to gain efficiencies in evaluation wherever feasible

The CTSA Consortium could benefit from negotiating contracts and licenses to access analytic services, software, data (e.g., bibliometric data), and other resources that all local CTSAs need to accomplish successful evaluation. In addition, CTSA program leaders should consider starting a Web site for sharing published reports, technical tips, procedures, reference documents, and the like.

The NIH needs to support the establishment of an identifiable evaluation entity within the new NCATS structure that oversees and manages CTSA evaluation activities

As part of its responsibilities, the evaluation entity within the NCATS structure should be charged with managing, guiding, facilitating, planning, and setting standards for national evaluations.

It should also be charged with conducting or overseeing national evaluation activities, including the coordination of the Evaluation Key Function Committee.

The CTSA evaluation community should recognize and draw on the wealth of resources it has among its own evaluation community and move to create a national virtual laboratory of evaluation

Local CTSAs should explore strategies for using internal capacity to further translational evaluation activities through coaching, mentoring, and joint exploratory activities. The existing CTSAs, with their varied profiles and approaches, provide the opportunity to create a national virtual laboratory of evaluation—a Web-based site in which knowledge can be synergized and innovation can be explored. While establishing such a laboratory will require opening doors and sharing operations, the potential benefit from bringing together a critical mass of clinical translational evaluators is considerable.

The Road Ahead

The CTSA initiative provides an historic opportunity to advance the field of translational science, but with this opportunity comes a significant responsibility. The initiative needs to show that it is well-implemented, efficiently managed, and demonstrably effective. Evaluation is key to achieving these goals. The NIH should be commended for the manner in which it has already incorporated evaluation into the CTSA initiative. However, the initial promise of evaluation and the ultimate realization of evaluation may not match. While the former clearly has been offered, the latter requires a strong and sustained commitment not only of resources but also, and perhaps more important, of human and organizational commitment. The recommendations presented in this document offer guidance for how the commitments can effectively be pursued in the context of the CTSA program. These commitments to and guidelines for evaluation will affect more than just our understanding of the CTSAs and how they function. They can meaningfully enhance our national ability to be accountable for what the CTSA program accomplishes. They can also add a critically important perspective to the emerging field of translational science and can serve as an historically important model for the multilevel evaluation of large scientific research initiatives.

Guidelines, by their very nature, are general. They are not meant to provide specific details about how they might best be addressed. The guidelines offered here are intended to be the foundation for a dialogue about how they can most effectively be translated into practice.

Conclusions

Evaluating translational science efforts is necessary for understanding the extent to which these initiatives are achieving their intended outcomes. As such, high quality evaluations must be an essential part of the CTSA program in order to provide the prospective and retrospective information necessary to direct its course and to assess the degree to which the program is accomplishing its goals. Currently, there is much discussion centered on CTSA evaluation questions, metrics, definitions, and procedures.

As members of the Evaluation Key Function Committee of the CTSA Consortium, we believe that despite 6 years of CTSA evaluation with some success, there is still considerable

need for a greater understanding of what evaluation is, why it is important, how it should be used, and how it fits into the CTSA landscape. In this paper, we set forth ideas that represent the collective thinking of numerous evaluators with considerable cumulative experience in struggling with the issues of how best to evaluate major scientific programs. These guidelines are designed to serve as a tool for education within the CTSA community by illuminating key issues and practices that should be considered during evaluation planning, implementation, and utilization. Additionally, these guidelines can serve as a basis for ongoing discussions about how the principles articulated in this paper can most effectively be translated into operational reality. While no single document can successfully address all questions and concerns about evaluation, we hope that the information in this paper will encourage the CTSA community to devise explicit policies regarding what evaluation is, what it is expected to be and how it can best be pursued. The anticipated result is that the knowledge generated by the CTSA evaluation processes will help inform decision making about how to best utilize the research resources that are available to achieve gains in public health.

Acknowledgments

We want to thank Meryl Sufian of the National Center for Advancing Translational Sciences, National Institutes of Health, Bethesda, Maryland, USA, and Joy A. Frechtling, Ph.D. of Westat, Rockville, Maryland, USA, for their outstanding leadership and collaboration on the original guidelines document. This project has been funded in whole or in part with Federal funds from the National Center for Research Resources and National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH), through the Clinical and Translational Science Awards Program (CTSA) by grants UL1 RR024996 and UL1 TR000457; UL1 RR024153 and UL1 TR000005; and UL1 RR031975 and UL1 TR000101. The manuscript was approved by the CTSA Consortium Publications Committee. The work was a collaborative effort of the CTSA Evaluation Key Function Committee and its National Evaluation Liaison Workgroup and was based in part on guidance from experts and documents in the field of evaluation. The leading professional association in the field, the AEA, has through its EPTF produced the document *An Evaluation Roadmap for a More Effective Federal Government*¹³ (the AEA Roadmap), which we have learned from and deliberately

emulated here. The AEA Roadmap has had and continues to have considerable influence on the U.S. federal government, including its legislative and executive branches.¹⁵ We received encouragement, consultation, and feedback from the AEA's EPTF in the formulation of our work, and while the opinions expressed here are those of the CTSA Evaluation Key Function Committee's National Evaluation Liaison Workgroup members and cannot be attributed to the AEA or its EPTF, our perspectives were very much enlightened and informed through our dialogue with members of the EPTF and other leaders within the AEA. We wish to thank them for their support in this effort.

References

1. Patton MQ. *Utilization-Focused Evaluation*. 4th ed. London, England: Sage; 2008.
2. American Evaluation Association. *Guiding Principles for Evaluators*. <http://www.eval.org/publications/GuidingPrinciplesPrintable.asp>. Ratified July 2004. Accessed May 11, 2011.
3. Yarbrough DB, Shulha LM, Hopson RK, Caruthers FA. *The Program Evaluation Standards: A Guide for Evaluators and Evaluation Users*. 3rd ed. Thousand Oaks, CA: Sage; 2011.
4. Dougherty D, Conway PH. The "3T's" road map to transform US health care: the "how" of high-quality care. *JAMA*. 2008; 299(19): 2319–2321.
5. Khoury MJ, Gwinn M, Yoon PW, Dowling N, Moore CA, Bradley L. The continuum of translation research in genomic medicine: how can we accelerate the appropriate integration of human genome discoveries into health care and disease prevention? *Genet Med*. 2007; 9(10): 665–674.
6. Sung NS, Crowley WFJr, Genel M, Salber P, Sandy L, Sherwood LM, Johnson SB, Catanese V, Tilson H, Getz K, et al. Central challenges facing the national clinical research enterprise. *JAMA*. 2003; 289(10): 1278–1287.
7. Trochim W, Kane C, Graham MJ, Pincus HA. Evaluating translational research: a process marker model. *Clin Transl Sci*. 2011; 4(3): 153–162.
8. Westfall JM, Mold J, Fagnan L. Practice-based research—"Blue highways" on the NIH roadmap. *JAMA*. 2007; 297(4): 403–406.
9. Rubio DM, Del Junco DJ, Bhore R, Lindsell CJ, Oster RA, Wittkowski KM, Welty LJ, Li YJ, Demets D. Biostatistics, Epidemiology, and Research Design (BERD) Key Function Committee of the Clinical and Translational Science Awards (CTSA) Consortium. Evaluation metrics for biostatistical and epidemiological collaborations. *Stat Med*. 2011; 30(23): 2767–2777.
10. Patton MQ. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York, NY: Guilford Press; 2011.
11. Williams B, Hummelbrunner R. *Systems Concepts in Action: A Practitioner's Toolkit*. Stanford, CA: Stanford University Press; 2010.
12. Williams B, Imam I, eds. *Systems Concepts in Evaluation: An Expert Anthology*. Point Reyes, CA: EdgePress; 2006.
13. American Evaluation Association. *An Evaluation Roadmap for a More Effective Government*. <http://www.eval.org/EPTF/aea10.roadmap.101910.pdf>. September 2010. Accessed May 11, 2011.
14. Cooks LJ, Mark MM, Trochim WM. Evaluation policy and evaluation practice: where do we go from here? *New Directions for Evaluation*. 2009; 123: 103–109.
15. American Evaluation Association. External Citations and Use of Evaluation Policy Task Force Work. <http://www.eval.org/EPTF/citations.asp>. Cited September 2011. Accessed September 7, 2011.