Program Evaluation for Managers

Primer

Management Decision and Research Center
Health Services Research and Development Service
Office of Research and Development
Department of Veterans Affairs

In collaboration with
Association for Health Services Research

The Health Services Research and Development Service (HSR&D) is a program within the Veterans Health Administration's Office of Research and Development. HSR&D provides expertise in health services research, a field that examines the effects of organization, financing and management on a wide range of problems in health care delivery — quality of care, access, cost and patient outcomes. Its programs span the continuum of health care research and delivery, from basic research to the dissemination of research results, and ultimately to the application of these findings to clinical, managerial and policy decisions.

Program Evaluation for Managers / by William Yeaton and Lois Camberg.
Boston, MA: Management Decision and Research Center; Washington, DC: VA Health Services Research and Development Service, in collaboration with Association for Health Services Research, 1997.
p. cm.
Includes bibliographical references. Primer.
DNLM: 1. Program Evaluation. 2. United States. Dept. of Veterans Affairs. Health Services Research and Development Service. 3. Association for Health Services Research
UH463.P964 1997

Program Evaluation for Managers

William Yeaton, Ph.D.
Lois Camberg, Ph.D.

Purpose of primer series: to help bridge the gap between health services researchers, policy makers, managers and clinicians in an effort to improve the quality and cost effectiveness of health care for veterans. The primer series is part of a larger set of dissemination initiatives developed by VHA's Office of Research and Development through its Management Decision and Research Center and in collaboration with the Association for Health Services Research.

Purpose of *Program Evaluation for Managers:* to introduce the purposes and general approaches to program evaluation. The primer provides a basic framework for understanding program evaluation and describes the potential of program evaluation as a management tool in health care settings, particularly in VA.

Suggested audience: professionals involved in health care decision making, including managers working in administration, clinical care, quality management and strategic planning at VA headquarters, Veterans Integrated Service Networks and within VA facilities.

Suggested uses: ■ individual study, ■ management training programs in VA networks and medical centers, ■ resource for strategic planning, ■ resource for quality management, ■ continuing medical education courses and other medical and health professional training programs.

August 1997
Boston, Massachusetts

Table of Contents

Contributors . . . . . *iii*

Preface . . . . . *v*

Introduction . . . . . *vii*

What is program evaluation? *1*

Why is program evaluation important? *1*

Can all programs be evaluated? *2*

When and how often should program evaluation be done? *2*

What resources are needed to conduct a program evaluation? *3*

What are the steps in planning a program evaluation? *4*

- *Step 1 — Explicitly state decision-maker(s), purpose(s) and potential applications, and formulate evaluation question(s)* *4*
- *Step 2 — Conduct a pre-evaluation assessment* *5*
- *Step 3 — Review the literature and contact administrative peers* *6*
- *Step 4 — Determine the evaluation design* *6*
- *Step 5 — Determine the interim communication process, scope of the final report and ultimate dissemination strategy* *6*

What are the potential limitations managers should be aware of regarding program evaluation? *6*

Concluding remarks *9*

Appendix A: Line items in sample program evaluation budget *11*

Appendix B: How are research designs classified? *13*

Appendix C: What organizations perform program evaluation? *15*

Appendix D: What reading materials are recommended? *19*

Contributors

William Yeaton, Ph.D., is a consultant based in Tallahassee, Florida. Dr. Yeaton worked for 12 years at the University of Michigan's Institute for Social Research (ISR), where his research focused on methodological issues in health and medicine. He also taught research and evaluation methods in health policy in the School of Public Health and continues to teach evaluation research design at ISR's Summer Institute.

Lois Camberg, Ph.D., is a health research scientist in the Health Services Research and Development Program at Brockton West Roxbury VA Medical Center, and an instructor at the Harvard University Medical School. She has worked on many federal and state program evaluations, taught workshops on program evaluation and coordinated the Management Consultation Program of VA's Management Decision and Research Center, which develops and manages program evaluations.

Acknowledgments. Special thanks go to Ann M. Hendricks, Ph.D., Center for Health Quality, Outcomes and Economic Research, Edith Nourse Rogers Memorial Veterans Affairs Hospital (Bedford, MA) and Chris Vogt, Ph.D., Arizona State University, for their contributions in developing the primer. Appreciation also is extended to colleagues throughout VA for their thoughtful comments in reviewing an earlier draft of the primer: Richard Adelson, D.D.S., Minneapolis Education Center; Joseph Bates, M.D., Little Rock VA Medical Center; Mike Berube, R.N., Little Rock VA Medical Center; Janet A. Deneselya, Ph.D., M.P.H., Butler VA Medical Center (Pennsylvania); Virginia Dodds, R.N., Baltimore VA Medical Center; Caryl Z. Lee, Veterans Integrated Service Network (Ann Arbor); Greg R. Neuner, M.B.A., Veterans Health Administration; John O'Neill, M.S.W., Boston VA Medical Center; John Sverha, Rates and Analysis Section, Medical Care Cost Recovery Program, Veterans Health Administration; and Sherry Treiber, VA Upstate New York Healthcare.

Preface

Policy makers, clinical managers and administrators make decisions based on intuition, common sense and information. Some decisions may be supportable by intuition or common sense alone, but complex, paradigm-changing decisions, or those with significant financial ramifications, should be guided by pertinent, accurate information or data.

We learn about the quality of a decision by observing its results. We learn most about the quality of a decision if we deliberately, systematically assess its impact or outcomes. Quantitative or qualitative data enable us to monitor and assess the "correctness" of our original decisions and help identify strategies for improvement. Enter the concept of program evaluation.

Unfortunately, the timing of decisions and the availability of quality data to optimize decision making are frequently not synchronized. Ideally, initial program development should build in monitoring systems to capture near-term and long-term data to provide information about how well the program is meeting its goals. Data inform and improve decisions. However, before any data are collected, the purpose(s) for the data collection must be clear. How are the data to be used to improve program management?

The scope of the program evaluation is determined by several factors — the importance of the evaluation questions, the quality of the data available, the complexity of evaluation questions and the availability of resources. The quality of the evaluation design determines the value of the program evaluation. While program evaluation is often done on the "quick and dirty," a deliberately planned and executed program evaluation is most likely to be useful to managers.

Although written by researchers, the primer is intended for use by non-researchers. It provides an introduction to the scope, methods, resources and limitations of the program evaluation process. Primer appendices provide further information, including bibliographic information and a list of other organizations conducting program evaluations.

John R. Feussner, M.D.
Chief Research and Development Officer
Veterans Health Administration

Introduction

Program evaluation is an important bridge between research and management. In this era of rapid change and emphasis on innovation and decentralized, effective and efficient management, program evaluation is an increasingly important and valued tool.

Program evaluations span a continuum from a simplistic counting to highly complex experimental design. Not all decisions require sophisticated research evaluations. For some decisions, the most appropriate evaluation is a simple tabulation or interview question. If the situation warrants research expertise to assure the quality of program evaluation, however, that expertise should be obtained.

When contemplating program evaluation, it's critical to clearly articulate what specific decisions will be based on the evaluation findings and who will be making them. What information is needed, why and by whom? A key question, for example, might be "What magnitude of what effect(s) over what time period would the decision maker consider appropriate evidence for program success?"

Next, evaluation planners need to determine the feasibility of the evaluation and if it's feasible, determine what type of program evaluation would be appropriate in view of the relative advantages and disadvantages of various approaches.

If a program evaluation involves sophisticated research techniques, is expected to yield findings generalizable beyond an individual facility, might contribute to the knowledge base of health services research *and* if time permits, the evaluation research team might develop a research protocol for peer review and potential funding.

This primer should assist those planning program evaluations of all types.

Shirley Meehan, M.B.A., Ph.D.
Acting Director, Health Services Research and Development
Veterans Health Administration

What is program evaluation?

Program evaluation is a method used to provide specific information about a clinical or administrative initiative's activities, outcomes, costs and effectiveness in meeting its goals. A "program" is a set of activities developed to accomplish one or more goals, as in a management information system or ambulatory surgery program.

Program evaluations vary in the technical complexity and sophistication of their research designs. Not all evaluations involve complex techniques or require specially trained research teams. Some can be done by managers who are not schooled in research methodologies, while others require the expertise of trained evaluators.

Why is program evaluation important?

As new programs are considered, evaluations from similar programs can assist with the decision of whether to implement the program and, if so, how to adapt it. As programs mature, occasional or ongoing evaluation can provide managers with information about adjustments that may be needed.

Ideally, program evaluation is used to determine if an initiative is meeting its stated objectives and, if relevant, the goals of the larger organization. Program evaluation also can produce specific information that managers can use to adjust or fundamentally change the structure and processes of the program to improve its outcomes.

Evidence from program evaluation is only one of many factors that influences decisions about a program.

■ Program evaluation can support short-term decision making about whether and how to modify an existing program:

- ❑ Does a specified information system work in meeting the information needs of clinical and financial decision makers?
- ❑ Is a hospital's smoking cessation program effective for specific ethnic populations?

■ Program evaluation also supports long-term decision making related to an organization's strategic plan:

- ❑ Is it likely that a service line organizational design being piloted in one VISN facility will be effective in all of its facilities?
- ❑ What impact does care delivered by subcontracted providers have on quality, cost and access as compared to care delivered in-house?

Evidence from program evaluation is only one of many factors that influences decisions about a program. Other factors, such as political ideology, the mix and influence of particular stakeholder groups and the availability of financial resources may also impact programmatic decision making.

Can all programs be evaluated?

In theory, all programs can be evaluated although any number of barriers can undermine the utility of an evaluation's findings. For example, politically unfavorable climates within and outside the organization can impede conducting a program evaluation. Or it may be imprudent to conduct an evaluation if the results are produceable only at great cost (relative to potential benefit) or are likely to be inferentially weak. Occasionally, time constraints may make a meaningful evaluation infeasible.

The decision to conduct a program evaluation is often reached during an evaluation feasibility study. (See discussion of pre-evaluation assessment on page 5.)

When and how often should program evaluation be done?

In an ideal world, program evaluation should be an ongoing process. However, scarce resources and political pressures often drive when and how often an evaluation is conducted. At a minimum, the decision about whether to conduct an evaluation should be considered at three times during a program's lifespan:

In an ideal world, program evaluation should be an ongoing process.

■ **When the program is being designed and first implemented.** For example, one could develop an evaluation design to examine the merits of implementing a nurse-supported telephone line by comparing outcomes of those having access to the telephone service with persons who did not have access. If just the persons with access were observed, the evaluators would only be able to describe the number of calls they made and their characteristics and outcomes.

■ **When the program or its environment is scheduled to change in some significant way.** Using the example above, it might be important to determine the impact on access and health outcomes if the size of the nursing staff or the number of nurse-supported line hours needed to be reduced.

■ **When alternative programs show promise for achieving better results (outcomes or cost) or when the effectiveness of the program is being questioned.** For example, the evaluator may need to consider the costs and benefits of using medical residents instead of nurses on the telephone support lines. A program evaluation could provide this information.

In addition, periodic evaluations are sometimes required during a program's lifespan (e.g., by the U.S. Congress).

What resources are needed to conduct a program evaluation?

Four resources are needed for an evaluation: an evaluation team, data, time and funding.

Four resources are needed for an evaluation: an evaluation team, data, time and funding.

■ **Evaluation team.** Not all evaluations require the expertise of trained researchers. Some can be done by administrators who are not schooled in research methodologies. For example, suppose a hospital is interested in evaluating whether it should start an evening clinic to increase working veterans' access to VA services. First, administrators might obtain mailing lists from Veterans Services Organizations and survey members to determine their interest in an evening clinic and their intention to use it. If members show enough interest, administrators could advertise the clinic's availability and conduct a few clinics to determine if patients would come. These efforts would not require highly sophisticated research skills and would provide adequate answers to the evaluation question.

The same administrative group could test, in the same way, the relative effectiveness of holding specialty or general clinics — or consider which week night(s) to implement them. Administrators also could distinguish the demographic and clinical characteristics of patients who attend evening clinics from those who attend daytime clinics, which might be useful in tailoring services. In contrast, expert evaluators would be needed if the hospital wanted to compare outcomes of persons treated in day clinics — controlling for the type and severity of illness presented — to those treated in evening clinics.

In its most technical form, program evaluation is a complex process requiring the technical skills and expertise of individuals from a wide range of research disciplines, including evaluation, statistics, organizational behavior, economics, political science, epidemiology, outcomes measurement and technology assessment. The exact makeup of the team will depend on the program being evaluated. Generally, the evaluation is overseen by a principal investigator, with the day-to-day activities managed by one or more mid-level coordinators.

Evaluation team members can be recruited from within the organization, outside the organization or a combination of both. Outside evaluators typically would need to be on-site intermittently throughout the evaluation period.

Consider internal evaluation teams when:

- ❑ The use of insiders could be advantageous because of their intimate knowledge of the program and its operation, if they have personal relationships with program staff that may facilitate more accurate, complete and efficient data collection.
- ❑ Strong resistance from program staff is anticipated, if there are financial or tight time constraints, or if in-house staff have the requisite knowledge, skills and objectivity to conduct the evalua-

tion.

Consider external evaluation teams when:

- ❑ Outside evaluators would provide a needed level of objectivity or new perspective that might be missed by internal evaluators.
- ❑ In-house staff do not have the essential skills and/or time needed to conduct an evaluation.
- ❑ The political climate requires an outside, "unbiased" assessment.

Note: When recruiting an outside evaluator, it is important to seek persons who have worked with similar organizations and who can describe an evaluation design that addresses the types of questions the program wants answered within the funding constraints. Soliciting multiple proposals from different evaluators should increase the chances of obtaining the highest quality of evaluation.

Program evaluation requires data--be it quantitative . . . or qualitative.

■ **Data.** Program evaluation requires data — be it quantitative (e.g., number of clinic visits) or qualitative (e.g., opinions shared in a focus group). The quality of an evaluation is greatly improved if pre-program, baseline data exist. Baseline data collected over a long period of time improve program evaluators' ability to assess needs and to draw valid conclusions. These baseline data may be collected for the sole purpose of conducting the evaluation or they may be constructed from existing databases. In the latter case, particular care should be taken to ensure the reliability and validity of baseline data.

■ **Time.** The evaluation plan should identify precise starting dates, interim milestones and ending dates. Deadlines should be realistic, yet consider the need for timely results required to make policy and management decisions.

■ **Funding.** The costs to conduct a program evaluation are as varied as program evaluations. Generally speaking, if outside evaluators are hired or new data are required, the evaluation will be more expensive. All costs of conducting an evaluation should be explicitly stated. These costs should be separated from the costs of running the program. (See Appendix A for a listing of items in a sample budget.)

These four resources typically are reviewed in the pre-evaluation assessment. (See discussion of pre-evaluation assessment on page 5.)

What are the steps in planning a program evaluation?

There is no cookbook approach or single recipe for planning a program evaluation. However, there are a number of characteristics common to most program evaluations. They include the following steps, which are not necessarily listed in order.

■ **Step #1 — Explicitly state decision-maker(s), purpose and potential applications, and formulate evaluation question(s).** The more explicit the understanding of who is the end-user of the evaluation, i.e., the decision-maker, why it is being conducted, and how the findings might be used, the better the evaluation. Will the evaluation findings be for internal use by program staff? Or will the findings be forwarded to the VISN director or perhaps disseminated to other similar programs across VA?

There is no cookbook approach or single recipe for planning a program evaluation.

What action steps will be influenced by evaluation findings? Answers to these questions are needed at the onset.

Asking good evaluation questions is the primary challenge to conducting sound program evaluation. Often, more than one question is asked. Some questions are better than others. Well-constructed research questions are focused, clearly stated and "measurable."

There are four types of questions asked in program evaluations.

❑ ***Needs assessment questions*** seek to identify needs that are being met, and those that are not.

For example: Will eligible, non-users of VA care use the VA system if evening clinics were offered?

❑ ***Formative or process questions*** address whether program elements were implemented as initially planned and if the target population is actually receiving the intended services. Information gained from formative questions allows for mid-course corrections.

For example: Can an adequate number and proper mix of providers be made available to conduct evening clinics on a regular basis?

❑ ***Summative questions*** ask if the program is improving important participant outcomes such as improved health status and quality of life or if participants' outcomes are better than non-participants.' Often, information gained from summative questions is used to determine the overall effectiveness of a program and to address whether it should continue.

For example: Do patients attending evening clinics realize improved health status?

❑ ***Financial questions*** assess the program from a fiscal standpoint (e.g., Is the program more cost-effective than an alternative program?)

For example: Do revenues from working veterans who use evening services cover the cost of conducting evening clinics?

■ Step #2 — Conduct a pre-evaluation assessment.
To best judge the wisdom of going ahead with an evaluation, a pre-evaluation assessment gauges the potential quality and usefulness of an evaluation.

A pre-evaluation assessment has three purposes: A) to clarify the intent of the program; B) to describe the program as it currently exists; and C) to assess the strengths of potential evaluation designs within the constraints of program design, the political climate, staff resources, available data, time and funding.

In conducting the pre-evaluation assessment, the evaluator examines program documentation and interviews relevant managers, policy makers and interest group members.

Findings from this pre-evaluation assessment ultimately influence

An evaluation design is strong or weak based on the resultant capacity to draw valid inference.

whether and when the evaluation should be conducted.

■ **Step #3 — Review the literature and contact administrative peers.** A thorough review of the literature will determine the degree and conditions under which important preliminary questions already have been answered and the quality of these answers. VA medical librarians can help with Medline searches. This review process also may inform the choice of a relevant research design.

Not all findings and knowledge are available in published form; peers are an additional and important source of information.

■ **Step #4 — Determine the evaluation design.** Selecting an appropriate research design depends on the question(s) being asked in the evaluation. For some evaluation questions, a methodologically rigorous research design is critical because it can greatly enhance the ability to draw valid conclusions.

An evaluation design is strong or weak based on the resultant capacity to draw valid inference. (See Appendix B for an annotated listing of research designs.)

■ **Step #5 — Determine the interim communication process, scope of the final report and ultimate dissemination strategy.** It should be determined from the start whether there will be interim reports, whether the final evaluation report will contain specific recommendations or merely list findings, and how and to whom evaluation findings will be disseminated.

What are the potential limitations managers should be aware of regarding program evaluation?

As with any research, managers should be mindful of the potential limitations of program evaluation.

As with any research, managers should be mindful of the potential limitations of program evaluation, since these limitations might alter the findings.

Conditional conclusions. All social science research assumes some degree of uncertainty in its conclusions. Evaluation results are true under specific conditions, and are usually couched in probabilities rather than stated with certainty.

Timeliness. Sometimes decisions about a program need to be made quickly. Program evaluation takes time to plan, conduct, analyze and report. Decisions may have to be made using the best available evidence, which may or may not come from the most rigorous and systematic evaluation.

Bias. Biases are systematic errors that may inadvertently influence the design, implementation and analysis of an evaluation and produce inaccurate results and conclusions.

Corruption of measures. Measures selected to inform and shape policy are sometimes incomplete and inaccurate and can distort the very programs they are intended to characterize. For example, when programs designed to find jobs for ex-military personnel

are judged "successful" based solely on the number of job placements, program managers may have given more attention to the number of placements than to the quality or appropriateness of the placements.

Applicability. Evaluation findings from one program may or may not be applicable to another program. Even if programs have very similar objectives and modes of operation, other factors unique to one program (e.g., program staff) may make it difficult to apply results to another program.

Invalid and unreliable data. The quality and type of data available to the research team are crucial to making meaningful conclusions. For example, it is often necessary to determine if available data are accurate or whether new data need to be collected. If new data are collected, the accuracy, consistency and relevance of the measures also need to be carefully scrutinized.

Unintended effects. Program evaluators must be alert to anticipate and measure effects that are unintended, but occur nonetheless (e.g. increased costs occurring with a program found to increase quality).

Concluding remarks

Any system--in order to stay relevant, effective and efficient--needs to continually ask questions about the component parts that make up its entirety.

There is an oft-told story in the evaluation research community about a famous physicist who, after receiving the Nobel Prize for Physics, dedicated the award to his mother. In his acceptance speech, he recounted a childhood story about how his mother would greet him each day when he came home from school with the same question. Many years later, he came to realize that her question was quite unusual. She never asked, "What did you learn in school today?" Instead, her question was, "What new questions did you learn to ask today?"

Program evaluation involves the art of both asking and answering questions. Policy makers, managers and clinicians alike use program evaluation as a tool to assist them in making informed decisions on the objectives, implementation and progress of their programs.

Any system — in order to stay relevant, effective and efficient — needs to continually ask questions about the component parts that make up its entirety. VA is an example of a large, complex system challenged by the need for maintaining highly efficient and responsive health care in the face of multiple changes (e.g., changes in the populations it treats, clinical practices, new technologies, federal policies and local environments). Each person in the VA system can and should take responsibility for asking evaluative questions.

Some evaluation questions can be answered easily with available information (e.g., how many rooms can be adequately cleaned with current housekeeping staff?). Other questions will require more long-term investigation and sophisticated techniques (e.g., to what extent does a new organizational model result in fewer adverse hospital events than the former model?). Each question asked offers VHA the opportunity to use new information to adjust current circumstances or create new circumstances that keep the system vital, well-tuned and ever responsive to its mission of providing quality health care services to veterans.

Appendix A: Line items in sample program evaluation budget

Budget category and line item	Comments
Personnel: Principal investigator (PI) — has overall responsibility for the final product	PI and/or project manager (PM) will usually be involved in other tasks. Often other collaborators are also included to provide complementary expertise.
Project manager (PM) — has responsibility for implementing and quality control of the project	
Program assistant (PA) — has responsibility for form production, correspondence and travel arrangements	Sometimes will perform data entry (see below).
Data collector (DC) — has responsibility for obtaining the information necessary for the project.	Number and qualifications of DCs depend on amount of data to be collected and type of data (e.g., interview vs. record review).
Data manager (DM) — has responsibility for making sure that all of the information collected is accurate and appropriately prepared for analysis	Not all projects require separate personnel for DM, data entry clerk (DE) and analyst (A).
Data entry clerk (DE) — has responsibility for entering data from raw format into computer-readable format	
Analyst (A) — has responsibility for computer-generated analyses of data collected	
Travel For data/information collection at study sites and meetings with funding source	Varies by number and location of sites and amount of information to be collected.
Supplies/photocopy Office supplies, document retrieval and study forms	
Equipment	A single program evaluation rarely requires equipment; office equipment is usually provided by evaluation agency.
Consultants Persons outside the agency with required expertise	

Appendix B: How are research designs classified?

There are many evaluation research designs, some of which are inferentially stronger than others. All evaluations should strive for methodological rigor, though in some cases the reality of the program and the setting in which it operates or the time available for evaluation will reduce the number of design options (e.g., experimental designs may not be possible or appropriate). The skill of a good evaluator lies in developing the strongest design within real-world constraints.

The research designs listed below are *roughly* grouped according to methodological rigor. It should be noted that for some evaluation questions, a "weaker" approach may be the most appropriate even in the absence of time and financial constraints.

The more methodologically complex the designs, the more likely is the need for the evaluation team to include persons trained in research methodology.

Research designs *without* control groups are generally viewed as having a relatively weak methodology:

■ ***Social indicators approach*** - examines descriptive statistics aggregated across groups of individuals to reflect the current status of some medical or health-related phenomenon. To illustrate, in-hospital mortality rates following bypass graft surgery may be used as a proxy for the quality of cardiovascular care in that hospital.

■ ***Case study design*** - reports the outcomes of an intervention in a single person or single group of persons. This approach can be used to look either back in time or forward in time. There are no control groups. In the health and medical literature, this design is also referred to as a "single case report" or "anecdote." For example, one might ask if satisfaction levels of VA physicians have been impacted subsequent to changes in reimbursement patterns for diagnostic tests.

■ ***Correlational design*** - statistically estimates the magnitude of the relationship between pairs of variables. There are no control groups. For example, one might reasonably ask if there is a negative correlation between the number of surgeries of a particular type performed at a given VA hospital and the number of deaths from that surgery.

Research designs *with* control groups are viewed as having a relatively strong methodology:

■ ***Historical control trials*** - compares the outcomes of a group of people who did not receive a program during some earlier

time period with the outcomes of a group that received the program in a later but separate period of time. In the last two years, stroke patients who received a particular drug regimen to limit subsequent cognitive deficit could be compared to patients who did not receive this regimen prior to this recent two year period.

■ ***Case-control design*** - typically compares the characteristics and treatment histories of a group of people who have a particular disease (cases) to an otherwise similar group of persons without the disease (controls). The case-control design uses data that are collected after the occurrence of the disease. For example, veterans with a particular syndrome might be compared to an otherwise similar group of veterans without the syndrome for differences in exposure to agents that may reasonably cause the disease.

■ ***Time-series (before and after) design*** - compares multiple outcomes in the same person or group of persons before and after their participation in a program or intervention. A comparison group that did not receive the program may be used to enhance inferential quality. In the event a legislative mandate required a particular change within all the VA hospitals of a state, outcomes pre-and post-change could be compared. These results could be contrasted to those found in another state in which this program change was not mandated.

■ ***Observed control trials*** - compares, forward in time, a group of individuals participating in a program with a group having similar characteristics that are not participating. The mortality rate of a group of veterans who have received angioplasty might be compared to a similar veteran group who have received a standard drug protocol for their cardiovascular disease.

■ ***Randomized control trials*** - compares, forward over time, the outcomes of persons, randomly assigned to one or more treatment or program groups to the outcomes of persons randomly assigned to a control (non-treatment) group. (In the vernacular of medical research, such a study is called a “clinical trial”.) A new and promising course of treatment for post traumatic stress disorder could be compared to the “usual care” approach by examining outcomes in groups of persons who have been randomly assigned to these two conditions.

Appendix C: What organizations perform program evaluation?

Although not exhaustive, listed below are a number of organizations within and outside VA that conduct program evaluation.

Inside VA

■ ***Office of Research and Development***
VA Headquarters
John R. Feussner, M.D., Chief Research and Development Officer
phone and FTS: 202/273-8284
Fax: 202/273-6526
VA's Office of Research and Development focuses its efforts on the health and care of our nation's veterans. The office oversees the full range of medical research in VA including: multi-site cooperative studies, rehabilitation research and health services research. Program evaluation spans these research activities.

❑ **Health Services Research and Development Service (HSR&D)**
VA Headquarters
Shirley Meehan, M.B.A., Ph.D., Acting Director
phone and FTS: 202/273-8287
Fax: 202/273-9007
Within the Office of Research and Development, HSR&D provides expertise in health services research, a field that examines the effects of health care organization, financing and management on a wide range of delivery issues, including quality of care, access, cost and patient outcomes. These programs span the continuum of health care research and delivery, from research to the dissemination of research results, and, ultimately, to the application of these findings to clinical, managerial and policy decisions.

HSR&D's key operating units are its nine Field Programs and the Management Decision and Research Center (below). HSR&D Field Programs are Centers of Excellence in targeted focus areas. These nine Field Programs foster the integration of research and practice, linking the clinical aspects of patient care with administrative needs through a core of VA staff at selected medical centers. Each program develops its own research agenda and maintains affiliations with community institutions — schools of public health, university health administration programs and research institutes — to support its objectives.

▲ Management Decision and Research Center (MDRC)
Boston VAMC
Martin P. Charns, D.B.A., Director
phone: 617/278-4433
FTS: 700/839-4433

Fax: 617/278-4438
The MDRC's mission is to enhance the delivery of the highest quality health care by providing VA senior staff with consultation, technical assistance, management information and research findings. MDRC's four interdependent programs — management consultation, information dissemination, technology assessment, and management and organizational research — provide VA researchers and managers powerful tools for planning and decision making, helping them find solutions to a wide range of problems in health care delivery.

❍ Management Consultation Program
Management Decision and Research Center
Boston VAMC
Carol Van Deusen Lukas, Ed.D., Manager
phone: 617/232-9500 ext. 5685
FTS: 700/839-4433
The Management Consultant Program provides senior VA managers with timely information to assist them in policy and decision making. Organized to respond to specific requests from VA health care managers, the program assists them with program evaluation, organizational development, technical analysis and research design. The program consults with the Health Services Research and Development Field Programs to perform program evaluation for senior VA managers.

■ *National Center for Health Promotion*

Durham VAMC
Robert J. Sullivan, Jr., M.D., M.P.H., Director
Laurence G. Branch, Ph.D., Associate Director
phone: 919/416-5880 x222
FTS: 700/671-5880 x222
Fax: 919/416-5879
The purpose of the VA National Center for Health Promotion (NCHP) is to help veterans receive optimal care by encouraging health promotion and disease prevention activities in facilities operated by the Department of Veterans Affairs. The NCHP staff track developments in the field, and disseminate information to clinicians with guidance regarding recommended procedures. Information on performance is analyzed to discern the best approach to help veterans reach their maximum health potential. The NCHP conducts program evaluations of health promotion and disease prevention programs.

■ *Northeast Program Evaluation Center*

West Haven VAMC
Robert Rosenheck, M.D., Director
phone: 203/937-3850
FTS: 700/428-3850
Fax: 203/937-3433
The Northeast Program Evaluation Center (NEPEC) conducts evaluations of national VA programs including initiatives in the areas of homelessness, chronic mental illness and post traumatic stress disorder. NEPEC evaluations are comprehensive, longitudinal and cross-sectional studies of programs including descriptions of participants, process evaluation of the initiatives and outcome assessment. Many of the evaluations conducted by NEPEC include systematic cost-

effectiveness analysis.
Outside VA

■ ***University-based evaluation programs***
Many schools of public health and other university-based programs employ senior researchers with expertise in program evaluation. Contact the following association to identify university-based program evaluation programs:

❑ **Association of Schools of Public Health**
Wendy Katz
1660 L Street, N.W.
Suite 204
Washington, DC 20036
phone: 202/296-1099
Fax: 202/296-1252
email: info@asph.org
website: www.asph.org

■ ***General Accounting Office (GAO)***
phone: 202/512-6000
website: www.gao.gov
GAO is the investigative arm of Congress and is charged with examining all matters relating to the receipt and disbursement of public funds. GAO performs a variety of services, including audits and evaluations of government programs and activities. The majority of these reviews are made in response to specific congressional requests. The Office offers a range of products to communicate the results of its work. Product types include testimony, oral briefings and written reports. All of GAO's unclassified reports are available to the public. A large number of products specific to VA are available on-line or from the U.S. Accounting Office, P.O. Box 6015, Gaithersburg, MD 20884-6015

■ **National databases and Internet listservs**
These also provide good information on program evaluation conducted in the private and public sector.

❑ ***American Evaluation Association (AEA)***
EVALTALK LISTSERV
Stephen Hebbler, List Manager
phone: 205/348-6801
e-mail: eal@uva1vm.ua.edu
website: www.eval.org
EVALTALK is a publicly-available, open discussion list devoted to issues in the field of evaluation. Participants include academic, public and private-sector researchers interested in evaluation as well as managers and administrators who implement and utilize program evaluation. Issues frequently discussed on EVALTALK include the development of evaluation questions and strategies for prioritizing questions generated by program evaluations, the role of evaluators, methodological design choices and fallback options, data analysis, processes for reporting evaluations, use of evaluations and evaluation data to help change programs and the politics of evaluations.

❑ ***Association for Health Services Research (AHSR) and National Library of Medicine (NLM)***
Health Services Research Projects (HSRProj) database
Naomi Adelman, M.L.S., Program Director (AHSR)
phone: 202/223-2477
Fax: 202/835-8972
website: *www.ahsr.org*
HSRProj is a growing database of over 4,000 records describing health services research grants and contracts funded by the public and private sectors. The database provides information about current research before the research results are published. Users can access HSRProj through NLM's MEDLARS system and retrieve information on a broad array of topics. The database provides primary investigator contact information to facilitate linkages among researchers working within a topic area, allowing them to share methodological and other lessons learned. A compendium of abstracts of ongoing program evaluation projects is available from AHSR upon request.

❑ ***National Library of Medicine (NLM)***
Health Services/Technology Assessment Research (HealthSTAR) database
National Information Center on Health Services Research and Health Care Technology (NICHSR)
phone: 301/496-0176
Fax: 301/402-3193
e-mail: nichsr@nlm.nih.gov
website: *www.nlm.nih.gov/nichsr/nichsr.html*
HealthSTAR focuses on the clinical (emphasizing the evaluation of patient outcomes and the effectiveness of procedures, programs, products, services and processes) and the non-clinical (emphasizing health care administration, economics planning and policy) aspects of health care delivery. HealthSTAR includes relevant bibliographic records from MEDLINE from 1975 to the present and additional specially indexed records for journal articles compiled by the American Hospital Association on health care administration; monographs, technical reports and theses from the National Health Planning Information Center; and journal articles, technical and government reports, meeting papers and abstracts and books and book chapters on health services research, clinical practice guidelines and health care technology assessment reports.

Appendix D: What reading materials are recommended?

The field of program evaluation is rich and growing. The limited space available in this primer has provided an opportunity to tell an important but necessarily incomplete story. For the interested reader, the authors have provided a list of more in-depth, written resources. These books and articles are arranged by relevant section of the primer.

What is program evaluation?

■ Cronbach, L.J., and Associates. *Toward Reform of Program Evaluation*. San Francisco: Jossey-Bass, 1980.

■ Fink, A.R., and Kosekoff, J. *An Evaluation Primer*. Washington, D.C.: Capital Publications Inc., 1978.

■ Rossi, P.H., and Freeman, H.E. *Evaluation: A Systematic Approach* (5th Edition). Beverly Hills: Sage Publications, 1993

■ Sechrest, L., and Figueredo, A.J. "Program Evaluation." *Annual Review of Psychology*, 44:1-31, 1993.

■ Shadish, W.R., Cook, T.D., and Leviton, L.C. *Foundations of Program Evaluation: Theories of Practice*. Newbury Park, CA: Sage Publications, 1991.

Why is program evaluation important?

■ Bennett, C.A., and Lumsdaine, A.A., eds. *Evaluation and Experiment*. New York: Academic Press, 1975.

■ Bickman, L., ed. "Advances in program theory." *New Directions for Program Evaluation*, Vol. 47. San Francisco: Jossey-Bass, 1990.

■ Braverman, M.T., and Campbell, D.T. "Facilitating the development of health promotion programs: Recommendations for researchers and funders." In Braverman, M.T., ed. *Evaluating Health Promotion Programs, New Directions for Program Evaluation*, Vol. 43:5-18 San Francisco: Jossey-Bass, 1989.

■ Caplan, N. "A minimal set of conditions necessary for the utilization of social science knowledge in policy formulation at the national level." In Weiss, C.H., ed. *Using Social Research in Public Policy Making*. Lexington, MA: Lexington-Heath, 1977.

■ Ciarlo, J. A., ed. *Utilizing Evaluation*. Sage Research Program Series in Evaluation, Vol. 6. Beverly Hills: Sage Publications, 1981.

■ Fairley, W.B., and Mosteller, F. *Statistics and Public Policy*. Reading, MA: Addison-Wesley, 1977.

■ Stevens, W.F., and Tornatzky, L.G. "The dissemination of evaluation: An experiment." *Evaluation Review*, 4(3):339-354 1980.

■ Kizer, K.W. *Vision for Change: A Brief summary*, U.S. Department of Veteran Affairs, Veterans Health Administration, March 1995.

■ U.S. Congress, Office of Technology Assessment. *The Impact of Clinical Trials on Health Policy and Medical Practice: Background Paper*. Washington: OTA-BP-H-22, 1983.

Can all programs be evaluated?

■ Wholey, J.S. *Evaluation: Promise and Performance*. Washington: Urban Institute, 1979.

When and how often should program evaluation be done?

■ Campbell, D.T. "Considering the case against experimental evaluations of social innovations." *Administrative Science Quarterly*, 15(1):110-113, 1970.

■ Creps, L.B., Coffey, R.J., Warner, P.A., and McLatchey, K.D. "Integrating total quality management and quality assurance at the University of Michigan Medical Center." *Quality Review Bulletin*, August, 1992.

■ Laffel, G., and Blumenthal, D. "The case for using industrial quality management science in health care organizations." *Journal of the American Medical Association*, 262(20):2869-2873, 1989.

What resources are needed to conduct a program evaluation?

■ Leviton, L.C., Hegedus, A.M., and Kubrin, A., eds. "Evaluating AIDS prevention: Contributions of multiple disciplines". *New Directions for Program Evaluation*, Vol. 46. San Francisco: Jossey-Bass, 1990.

■ Love, A.J., ed. "Developing effective internal evaluation." *New Directions for Program Evaluation*, Vol. 20. San Francisco: Jossey-Bass, 1983.

■ Reiss, A.J., and Boruch, R.. "The program review team approach and multi-site experiments: The spouse assault replication program." In Turpin, R.S. and Sinacore J.M., eds. *Multi-site Evaluations, New Directions for Program Evaluation,* Vol. 50:33-44. San Francisco: Jossey-Bass, 1991.

What are the steps in planning a program evaluation?

Step #1 – Explicitly state decision-maker(s), purpose and potential applications, and formulate evaluation question(s)

■ Albritton, R.B. "Cost-benefits of measles eradication: Effects of a federal intervention." *Policy Analysis*, 4(1):1-22, 1978.

■ Lipsey, M.W. *Design Sensitivity. Statistical Power for Experimental Research.* Newbury Park: Sage Publications, 1990.

■ Riecken, H.W., and Boruch, R.F. *Social Experimentation: A Method for Planning and Evaluating Social Intervention*. New York: Academic Press, 1974.

Step #2 – Conduct a pre-evaluation assessment

■ Horst, P., Nay, J.N., and Wholey, J.S. "Program management and the federal evaluator." *Public Administration Review*, 34(4):300-308, 1974.

■ Light, R.J. "President's corner." *Evaluation Practice*, 7(3):87-91, 1986.

■ Wholey, J.S. *Evaluation and Effective Public Management*. Boston: Little and Brown, 1983.

Step #3 – Review the literature and contact administrative peers

■ Chelimsky, E., and Morra, L.G. "Evaluation synthesis for the legislative user." In Yeaton ,W.H., and Wortman, P. M., eds. *Issues in Data Synthesis. New Directions for Program Evaluation*, Vol. 24:75-89. San Francisco: Jossey-Bass, 1984.

■ Cooper, H.M., and Hedges, L.V., eds. *The Handbook of Research Synthesis*. New York: Russell Sage Foundation, 1994.

■ Cordray, D.S. "Strengthening causal interpretations of non-experimental data: The role of meta-analysis." *New Directions for Program Evaluation*, Vol. 60:59-96. San Francisco: Jossey-Bass, 1993.

■ Government Accounting Office, Program Evaluation and Methodology Division. *Cross-design Synthesis: A New Strategy for Medical Effectiveness Research*. Washington: Government Printing Office, # 92-18, 1992.

■ Hedrick, S.C., Koepsell, T.D., and Inui, T. "Meta-analysis of homecare effects on mortality and nursing home placement." *Medical Care*, 27(11):1015-1026, 1989.

■ Hunter, J.E., and Schmidt, F.L. *Methods of Meta-analysis*. Newbury Park: Sage, 1990

■ Light, R.J. "Six evaluation issues that synthesis can resolve better than single studies." In Yeaton, W.H., and Wortman, P.M., eds. *Issues in Data Synthesis: New Directions for Program Evaluation,* Vol. 24:57-73. San Francisco: Jossey-Bass 1984.

■ Yeaton, W.H., and Wortman, P.M. "Medical technology assessment: The evaluation of coronary artery bypass graft surgery using data synthesis techniques." *International Journal of Technology Assessment in Health Care*, 1(2):125-136, 1985.

Step #4 – Determine the evaluation design

■ Campbell, D.T., and Stanley, J.C. *Experimental and Quasi-experimental Designs for Research*. Chicago: Rand McNally, 1963.

■ Cook, T.D., and Campbell, D.T. *Quasi-experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally, 1979.

■ Freedman, D.A. "Statistical models and shoe leather." In Marsden, P. ed. *Sociological Methodology*, Vol. 21:291-313. Oxford: Basil Blackwell LTD, 1991.

■ General Accounting Office, Program Evaluation and Methodology Division. *Case Study Evaluations*. Transfer Paper 9, 1987.

■ Hill, A.B. "The environment and disease: Association or causation?" *Proceedings of the Royal Society of Medicine*, 58(May):295-300, 1965.

■ Horwitz, R.I., and Feinstein, A.R. "Methodologic standards and contradictory results in case-control research." *The American Journal of Medicine*, 66(4):556-564, 1979.
■ Kazdin, A.E. "Drawing valid inferences from case studies." *Journal of Consulting and Clinical Psychology*, 49(2):183-192, 1981.

■ Light, R.J., Singer, J.D., and Willett, J.B. *By Design*. Cambridge, MA: Harvard University Press, 1990.

■ Meier, P. "The biggest public health experiment ever: The 1954 field trial of the Salk poliomyelitis vaccine." In Tanur, J.M., et al., eds. *Statistics: A Guide to the Unknown*. Berkeley, CA: Holden, 1978.

■ Phillips, K.A., Luft, H.S., and Ritchie, J.L. "The association of hospital volumes of percutaneous transluminal coronary angioplasty with adverse outcomes, length of stay, and charges in California." *Medical Care*, 33(5):502-514, 1995.

■ Sachs, H., Chalmers, T.C., and Smith, H. "Randomized versus historical controls for clinical trials." *The American Journal of Medicine*, 72(2):233-240, 1982.

■ Shadish, W.R., Cook, T.D., and Houts, A.C. "Quasi-experimentation in a critical multiplist mode." In Trochim, W.M.K., ed. *Advances in Quasi-experimental Design and Analysis: New Directions for Program Evaluation*, Vol. 31:29-46. San Francisco: Jossey-Bass, 1986.

■ Yeaton, W.H., and Sechrest, L. "Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness." *Journal of Consulting and Clinical Psychology*, 49(2):156-167, 1981.

■ Yeaton, W.H., and Sechrest, L. "Use and misuse of no-difference findings in eliminating threats to validity." *Evaluation Review*, 10(6):836-852, 1987.

■ Yin, R.K. *Case Study Research: Design and Methods*, 2nd ed. Thousand Oaks: Sage Publications, 1994.

Step #5 – Determine the interim communication process, scope of the final report and ultimate dissemination strategy.

■ Lyons Morris, L., Taylor Fitz-Gibbon, C., and Freeman, M.E. *How to Communicate Evaluation Findings*. Newbury Park: Sage Publications, Inc., 1987

What are the potential limitations managers should be aware of regarding program evaluation?

■ Ginsberg, P.N. "The dysfunctional side effects of quantitative indicator production: Illustrations from mental health care." *Evaluation and Program Planning*, 7:1-12 1984.

■ Rabeneck, L. *Development of illness severity scales for HIV infection. Projects receiving Funding in Fiscal year 1994*. Veterans Administration HSR and D Document #39-32-009, p. 30, 1994.

■ Sechrest, L. "Approaches to ensuring quality of data and performance: Lessons for science?" In Jackson, D.N., and Rushton, J.P. eds. *Scientific Excellence: Origins and Assessment*. Beverly Hills: Newbury Park, 1987.

Fax us your comments!

To: MDRC

Fax: 617/278-4438

From: _____
name

_____
title

_____
address/facility

_____

_____

How will you use the program evaluation primer? ***(check all that apply)***

___ for my own education/information

___ to work with other staff members to increase understanding of program evaluation

___ as a meeting/conference/inservice training handout

___ other *(please specify)*_____

_____

_____

Please rate the amount of information provided ***(circle one)***

(1=not enough 5=just right) *1* *2* *3* *4* *5*

What is your overall rating of the primer? ***(circle one)***

(1=not helpful 5=very helpful) *1* *2* *3* *4* *5*

General comments _____

_____

_____

_____

_____

Suggestions for future primer topics_____

_____

_____

_____

P*rogram Evaluation for Managers* is available in electronic and printed formats. Additional copies may be obtained from the sources listed below.

Electronic copies (PDF format) can be downloaded from the VA home page on the world wide web. Point your browser to http://www.va.gov/mdrc.

Faxed copies:
From a fax machine or telephone, call
617/278-4492
Follow the voice menu system instructions to order the program evaluation primer. Requesters are advised that fax transmission can take up to 30 minutes.

Printed copies-contact:
Special Projects Office (152)
VA Medical Center
Perry Point, MD 21902
VA FTS System 700/956-5442
Commercial 410/642-1092
Fax 410/642-1095
E-mail: long.laurel@forum.va.gov

Other primers in the series include:

Primary Care in VA

Health Care Technology Assessment in VA

Using Outcomes to Improve Healthcare Decision Making