CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

September 17, 1999 / Vol. 48 / No. RR-11

MMWR™
MORBIDITY AND MORTALITY
WEEKLY REPORT

Recommendations and Reports

Inside: Continuing Education Examination

Framework for Program Evaluation in Public Health

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES
Centers for Disease Control and Prevention (CDC)
Atlanta, Georgia 30333

The *MMWR* series of publications is published by the Epidemiology Program Office, Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services, Atlanta, GA 30333.

SUGGESTED CITATION

Centers for Disease Control and Prevention. Framework for program evaluation in public health. MMWR 1999;48(No. RR-11):[inclusive page numbers].

Copies can be purchased from Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402-9325. Telephone: (202) 512-1800.

Contents

FOREWORD

Health improvement is what public health professionals strive to achieve. To reach this goal, we must devote our skill — and our will — to evaluating the effects of public health actions. As the targets of public health actions have expanded beyond infectious diseases to include chronic diseases, violence, emerging pathogens, threats of bioterrorism, and the social contexts that influence health disparities, the task of evaluation has become more complex. CDC developed the framework for program evaluation to ensure that amidst the complex transition in public health, we will remain accountable and committed to achieving measurable health outcomes.

By integrating the principles of this framework into all CDC program operations, we will stimulate innovation toward outcome improvement and be better positioned to detect program effects. More efficient and timely detection of these effects will enhance our ability to translate findings into practice. Guided by the steps and standards in the framework, our basic approach to program planning will also evolve. Findings from prevention research will lead to program plans that are clearer and more logical; stronger partnerships will allow collaborators to focus on achieving common goals; integrated information systems will support more systematic measurement; and lessons learned from evaluations will be used more effectively to guide changes in public health strategies.

Publication of this framework also emphasizes CDC's continuing commitment to improving overall community health. Because categorical strategies cannot succeed in isolation, public health professionals working across program areas must collaborate in evaluating their combined influence on health in the community. Only then will we be able to realize and demonstrate the success of our vision — healthy people in a healthy world through prevention.

Jeffrey P. Koplan, M.D., M.P.H.
Director, Centers for Disease Control
and Prevention
Administrator, Agency for
Toxic Substances and Disease Registry

The following CDC staff members prepared this report:

Robert L. Milstein, M.P.H.
Office of Program Planning and Evaluation
Office of the Director

Scott F. Wetterhall, M.D., M.P.H.
Chair, CDC Evaluation Working Group
Office of Program Planning and Evaluation
Office of the Director

in collaboration with
CDC Evaluation Working Group Members

Gregory M. Christenson, Ph.D.
Diane Dennis-Flagler
Division of Health Education and Promotion
Agency for Toxic Substances and Disease Registry

Jeffrey R. Harris, M.D.
Donna L. Higgins, Ph.D.
Kenneth A. Schachter, M.D., M.B.A.
Division of Prevention Research and Analytic Methods
Nancy F. Pegg, M.B.A.
Office of the Director
Epidemiology Program Office

Janet L. Collins, Ph.D., M.S.
Division of Adolescent and School Health
Diane O. Dunet, M.P.A.
Division of Cancer Prevention and Control
Aliki A. Pappas, M.P.H., M.S.W.
Division of Oral Health
National Center for Chronic Disease Prevention and Health Promotion

Alison E. Kelly, M.P.I.A.
Office of the Director
National Center for Environmental Health

Paul J. Placek, Ph.D.
Office of Data Standards, Program Development, and Extramural Programs
National Center for Health Statistics

Michael Hennessy, Ph.D., M.P.H.
Division of STD Prevention
Deborah L. Rugg, Ph.D.
Division of HIV/AIDS Prevention — Intervention, Research, and Support
National Center for HIV, STD, and TB Prevention

Additional CDC Contributors

Office of the Director: Lynda S. Doll, Ph.D., M.A.; Charles W. Gollmar; Richard A. Goodman, M.D., M.P.H.; Wilma G. Johnson, M.S.P.H.; Marguerite Pappaioanou, D.V.M., Ph.D., M.P.V.M.; David J. Sencer, M.D., M.P.H. (Retired); Dixie E. Snider, M.D., M.P.H.; Marjorie A. Speers, Ph.D.; Lisa R. Tylor; and Kelly O'Brien Yehl, M.P.A. (Washington, D.C.).

Agency for Toxic Substances and Disease Registry: Peter J. McCumiskey and Tim L. Tinker, Dr.P.H., M.P.H.

Epidemiology Program Office: Jeanne L. Alongi, M.P.H. (Public Health Prevention Service); Peter A. Briss, M.D.; Andrew L. Dannenberg, M.D., M.P.H.; Daniel B. Fishbein, M.D.; Dennis F. Jarvis, M.P.H.; Mark L. Messonnier, Ph.D., M.S; Bradford A. Myers; Raul A. Romaguera, D.M.D., M.P.H.; Steven B. Thacker, M.D., M.Sc.; Benedict I. Truman, M.D., M.P.H.; Katherine R. Turner, M.P.H. (Public Health Prevention Service); Jennifer L. Wiley, M.H.S.E. (Public Health Prevention Service); G. David Williamson, Ph.D.; and Stephanie Zaza, M.D., M.P.H.

National Center for Chronic Disease Prevention and Health Promotion: Cynthia M. Jorgensen, Dr.P.H.; Marshall W. Kreuter, Ph.D., M.P.H.; R. Brick Lancaster, M.A.; Imani Ma'at, Ed.D., Ed.M., M.C.P.; Elizabeth Majestic, M.S., M.P.H.; David V. McQueen, Sc.D., M.A.; Diane M. Narkunas, M.P.H.; Dearell R. Niemeyer, M.P.H.; and Lori B. de Ravello, M.P.H.

National Center for Environmental Health: Jami L. Fraze, M.S.Ed.; Joan L. Morrissey; William C. Parra, M.S.; Judith R. Qualters, Ph.D.; Michael J. Sage, M.P.H.; Joseph B. Smith; and Ronald R. Stoddard.

National Center for Health Statistics: Marjorie S. Greenberg, M.A. and Jennfier H. Madans, Ph.D.

National Center for HIV, STD, and TB Prevention: Huey-Tsyh Chen, Ph.D.; Janet C. Cleveland, M.S.; Holly J. Dixon; Janice P. Hiland, M.A.; Richard A. Jenkins, Ph.D.; Jill K. Leslie; Mark N. Lobato, M.D.; Kathleen M. MacQueen, Ph.D., M.P.H.; and Noreen L. Qualls, Dr.P.H., M.S.P.H.

National Center for Injury Prevention and Control: Christine M. Branche, Ph.D.; Linda L. Dahlberg, Ph.D.; and David A. Sleet, Ph.D., M.A.

National Immunization Program: Susan Y. Chu, Ph.D., M.S.P.H. and Lance E. Rodewald, M.D.

National Institute for Occupational Safety and Health: Linda M. Goldenhar, Ph.D. and Travis Kubale, M.S.W.

Public Health Practice Program Office: Patricia Drehobl, M.P.H.; Michael T. Hatcher, M.P.H.; Cheryl L Scott, M.D., M.P.H.; Catherine B. Shoemaker, M.Ed.; Brian K. Siegmund, M.S.Ed., M.S.; and Pomeroy Sinnock, Ph.D.

Consultants and Contributors

Suzanne R. Adair, Ph.D., M.A., Texas Department of Health, Austin, Texas; Mary Eden Avery, M.A., American Cancer Society, Atlanta, Georgia; Ronald Bialek, M.P.P., Public Health Foundation, Washington, D.C.; Leonard Bickman, Ph.D., Vanderbilt University, Nashville, Tennessee; Thomas J. Chapel, M.S., M.B.A., Macro International, Atlanta, Georgia; Don Compton, Ph.D., American Cancer Society, Atlanta, Georgia; Ross F. Conner, Ph.D., University of California Irvine, Irvine, California; David A. Cotton, Ph.D., M.P.H., Macro International, Atlanta, Georgia; Bruce Davidson, M.D., M.P.H., National Tuberculosis Controllers Association, Philadelphia, Pennsylvania; Mary V. Davis, Dr.P.H., Association of Teachers of Preventive Medicine, Washington, D.C.; William W. Dyal, DeKalb County Board of Health, Decatur, Georgia; Stephen B. Fawcett, Ph.D., University of Kansas, Lawrence, Kansas; Jane Ford, National Association of City and County Health Officers, Lincoln, Nebraska; Nicholas Freudenberg, Dr.P.H., M.P.H., City University of New York, New York, New York; Jean Gearing, Ph.D., M.P.H., DeKalb County Board of Health, Decatur, Georgia; Kristine Gebbe, Dr.P.H., Columbia University, New York, New York; David N. Gillmore, M.Ed., University of Texas School of Public Health, Houston, Texas; Rebecca M. Glover Kudon, M.S.P.H., American Cancer Society, Atlanta, Georgia; Lynne E. Greabell, M.A.A., National Association of State and Territorial AIDS Directors, Washington, D.C.; Susan R. Griffin, M.P.A., Independent Consultant, Austin, Texas; Sharon Lee Hammond, Ph.D., M.A., Westat, Inc., Atlanta, Georgia; Anne C. Haddix, Ph.D., Rollins School of Pubic Health, Atlanta, Georgia; Susan E. Hassig, Dr.P.H., Tulane School of Public Health and Tropical Medicine, New Orleans, Louisiana; Gary T. Henry, Ph.D., M.A., Georgia State University, Atlanta, Georgia; James C. Hersey, Ph.D., M.S., M.A., Research Triangle Institute, Research Triangle Park, North Carolina; Richard E. Hoffman, M.D., M.P.H., Council of State and Territorial Epidemiologists, Denver, Colorado; Robert C. Hornik, Ph.D., M.A., Annenberg School of Communication, Philadelphia, Pennsylvania; Eric Juzenas, American Public Health Association, Washington, D.C.; Mark R. Keegan, M.B.A., Association of State and Territorial Health Officers, Denver, Colorado; Jeffrey J. Koshel, M.A., Department of Health and Human Services, Washington, D.C.; Amy K. Lewis, M.P.H., North Carolina Department of Health and Human Services, Raleigh, North Carolina; Jennifer M. Lewis, M.Ed., Association of Schools of Public Health, Chapel Hill, North Carolina; Hardy D. Loe, Jr., M.D., M.P.H., University of Texas School of Public Health, Houston, Texas; Anna Marsh, Substance Abuse and Mental Health Service Administration, Washington, D.C.; Pamela Mathison, M.A., Texas Department of Health, Austin, Texas; Dennis McBride, Ph.D., University of Washington, Seattle, Washington; Kathleen R. Miner, Ph.D., M.P.H., Rollins School of Public Health, Atlanta, Georgia; April J. Montgomery, M.H.A., Colorado Department of Health, Denver, Colorado; Genevieve A. Nagy, University of Kansas, Lawrence, Kansas; Dennis P. Murphy, M.A., National Coalition of STD Directors, Albany, New York; Patricia P. Nichols, M.S., Michigan Department of Education, Lansing, Michigan; Mary Odell Butler, Ph.D., M.A., Battelle, Arlington, Virginia; Carol Pitts, Department of Health and Human Services, Washington, D.C.; Hallie Preskill, Ph.D., University of New Mexico, Albuquerque, New Mexico; Carol Roddy, Public Health Service, Washington, D.C.; Ken Duane Runkle, M.A., Illinois Department of Public Health, Springfield, Illinois; James R. Sanders, Ph.D., M.S.Ed., Western Michigan University, Kalamazoo, Michigan; Linda M. Scarpetta, M.P.H.,

Michigan Department of Community Health, Lansing, Michigan; Mark V. Schrader, M.P.A., Georgia Human Resources Department, Atlanta, Georgia; Leonard M. Schwartz, M.A., Michigan Department of Community Health, Lansing, Michigan; Harvey A. Schwartz, Ph.D., M.B.A., M.S., M.Ph., Agency for Health Care Policy and Research, Washington, D.C.; William A. Smith, Ed.D., Academy for Educational Development, Washington, D.C.; Edith M. Sternberg, M.P.H., Illinois Department of Health, Springfield, Illinois; Susan Toal, M.P.H., Independent Consultant, Atlanta, Georgia; Juliana van Olphen, M.P.H., University of Michigan, Ann Arbor, Michigan; Paul J. Wiesner, M.D., DeKalb County Board of Health, Decatur, Georgia; Joseph S. Wholey, Ph.D., University of Southern California, Los Angeles, California; Alice Wojciak, Association of Schools of Public Health, Washington, D.C.; and Robert K. Yin, Ph.D., Cosmos Corporation, Bethesda, Maryland.

Framework for Program Evaluation in Public Health

Summary

Effective program evaluation is a systematic way to improve and account for public health actions by involving procedures that are useful, feasible, ethical, and accurate. The framework guides public health professionals in their use of program evaluation. It is a practical, nonprescriptive tool, designed to summarize and organize essential elements of program evaluation. The framework comprises steps in program evaluation practice and standards for effective program evaluation. Adhering to the steps and standards of this framework will allow an understanding of each program's context and will improve how program evaluations are conceived and conducted. Furthermore, the framework encourages an approach to evaluation that is integrated with routine program operations. The emphasis is on practical, ongoing evaluation strategies that involve all program stakeholders, not just evaluation experts. Understanding and applying the elements of this framework can be a driving force for planning effective public health strategies, improving existing programs, and demonstrating the results of resource investments.

INTRODUCTION

Program evaluation is an essential organizational practice in public health (*1*); however, it is not practiced consistently across program areas, nor is it sufficiently well-integrated into the day-to-day management of most programs. Program evaluation is also necessary for fulfilling CDC's operating principles for guiding public health activities, which include a) using science as a basis for decision-making and public health action; b) expanding the quest for social equity through public health action; c) performing effectively as a service agency; d) making efforts outcome-oriented; and e) being accountable (*2*). These operating principles imply several ways to improve how public health activities are planned and managed. They underscore the need for programs to develop clear plans, inclusive partnerships, and feedback systems that allow learning and ongoing improvement to occur. One way to ensure that new and existing programs honor these principles is for each program to conduct routine, practical evaluations that provide information for management and improve program effectiveness.

This report presents a framework for understanding program evaluation and facilitating integration of evaluation throughout the public health system. The purposes of this report are to

- summarize the essential elements of program evaluation;
- provide a framework for conducting effective program evaluations;
- clarify the steps in program evaluation;
- review standards for effective program evaluation; and

- address misconceptions regarding the purposes and methods of program evaluation.

BACKGROUND

Evaluation has been defined as systematic investigation of the merit, worth, or significance of an object (*3,4*). During the past three decades, the practice of evaluation has evolved as a discipline with new definitions, methods, approaches, and applications to diverse subjects and settings (*4–7*). Despite these refinements, a basic organizational framework for program evaluation in public health practice had not been developed. In May 1997, the CDC Director and executive staff recognized the need for such a framework and the need to combine evaluation with program management. Further, the need for evaluation studies that demonstrate the relationship between program activities and prevention effectiveness was emphasized. CDC convened an Evaluation Working Group, charged with developing a framework that summarizes and organizes the basic elements of program evaluation.

Procedures for Developing the Framework

The Evaluation Working Group, with representatives from throughout CDC and in collaboration with state and local health officials, sought input from eight reference groups during its year-long information-gathering phase. Contributors included

- evaluation experts,
- public health program managers and staff,
- state and local public health officials,
- nonfederal public health program directors,
- public health organization representatives and teachers,
- community-based researchers,
- U.S. Public Health Service (PHS) agency representatives, and
- CDC staff.

In February 1998, the Working Group sponsored the Workshop To Develop a Framework for Evaluation in Public Health Practice. Approximately 90 representatives participated. In addition, the working group conducted interviews with approximately 250 persons, reviewed published and unpublished evaluation reports, consulted with stakeholders of various programs to apply the framework, and maintained a website to disseminate documents and receive comments. In October 1998, a national distance-learning course featuring the framework was also conducted through CDC's Public Health Training Network (*8*). The audience included approximately 10,000 professionals. These information-sharing strategies provided the working group numerous opportunities for testing and refining the framework with public health practitioners.

Defining Key Concepts

Throughout this report, the term *program* is used to describe the object of evaluation, which could be any organized public health action. This definition is deliberately broad because the framework can be applied to almost any organized public health activity, including direct service interventions, community mobilization efforts, research initiatives, surveillance systems, policy development activities, outbreak investigations, laboratory diagnostics, communication campaigns, infrastructure-building projects, training and educational services, and administrative systems. The additional terms defined in this report were chosen to establish a common evaluation vocabulary for public health professionals.

Integrating Evaluation with Routine Program Practice

Evaluation can be tied to routine program operations when the emphasis is on practical, ongoing evaluation that involves all program staff and stakeholders, not just evaluation experts. The practice of evaluation complements program management by gathering necessary information for improving and accounting for program effectiveness. Public health professionals routinely have used evaluation processes when answering questions from concerned persons, consulting partners, making judgments based on feedback, and refining program operations (*9*). These evaluation processes, though informal, are adequate for ongoing program assessment to guide small changes in program functions and objectives. However, when the stakes of potential decisions or program changes increase (e.g., when deciding what services to offer in a national health promotion program), employing evaluation procedures that are explicit, formal, and justifiable becomes important (*10*).

ASSIGNING VALUE TO PROGRAM ACTIVITIES

Questions regarding values, in contrast with those regarding facts, generally involve three interrelated issues: merit (i.e., quality), worth (i.e., cost-effectiveness), and significance (i.e., importance) (*3*). If a program is judged to be of merit, other questions might arise regarding whether the program is worth its cost. Also, questions can arise regarding whether even valuable programs contribute important differences. Assigning value and making judgments regarding a program on the basis of evidence requires answering the following questions (*3,4,11*):

- What will be evaluated? (That is, what is the program and in what context does it exist?)
- What aspects of the program will be considered when judging program performance?
- What standards (i.e., type or level of performance) must be reached for the program to be considered successful?
- What evidence will be used to indicate how the program has performed?
- What conclusions regarding program performance are justified by comparing the available evidence to the selected standards?

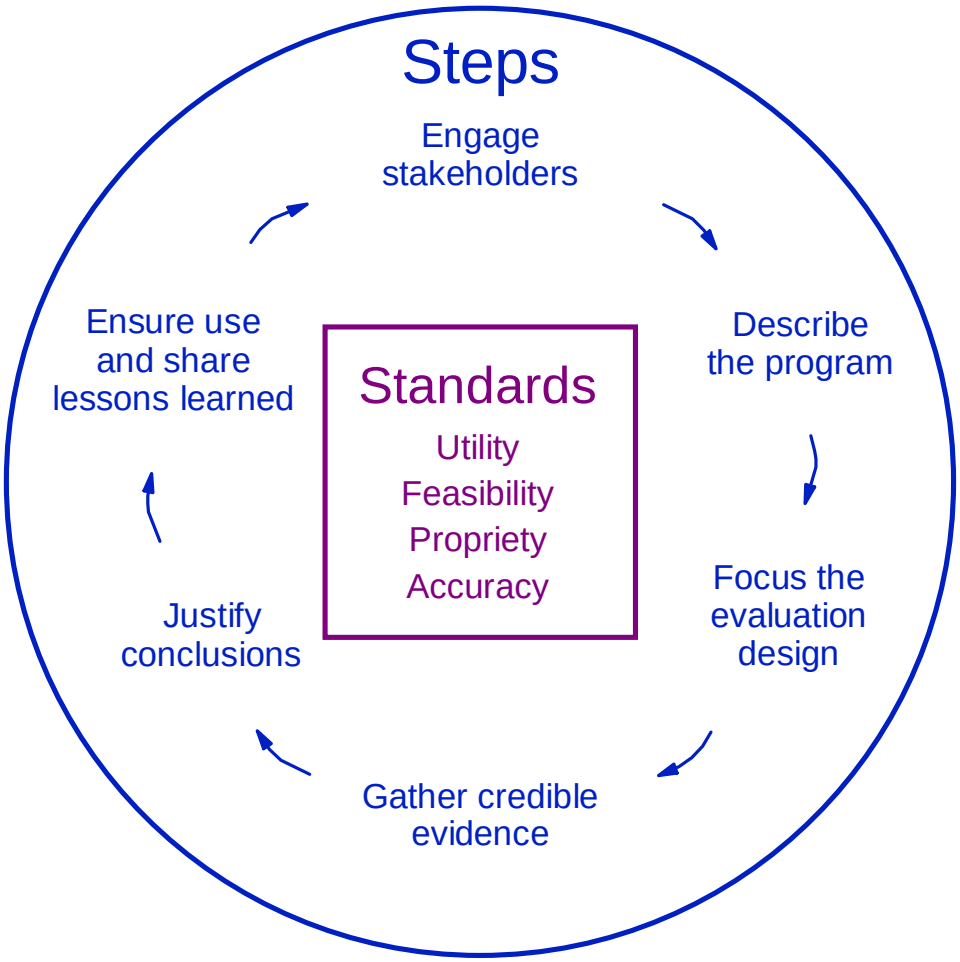
- How will the lessons learned from the inquiry be used to improve public health effectiveness?

These questions should be addressed at the beginning of a program and revisited throughout its implementation. The framework described in this report provides a systematic approach for answering these questions.

FRAMEWORK FOR PROGRAM EVALUATION IN PUBLIC HEALTH

Effective program evaluation is a systematic way to improve and account for public health actions by involving procedures that are useful, feasible, ethical, and accurate. The recommended framework was developed to guide public health professionals in using program evaluation. It is a practical, nonprescriptive tool, designed to summarize and organize the essential elements of program evaluation. The framework comprises steps in evaluation practice and standards for effective evaluation (Figure 1).

FIGURE 1. Recommended framework for program evaluation



The framework is composed of six steps that must be taken in any evaluation. They are starting points for tailoring an evaluation to a particular public health effort at a particular time. Because the steps are all interdependent, they might be encountered in a nonlinear sequence; however, an order exists for fulfilling each — earlier steps provide the foundation for subsequent progress. Thus, decisions regarding how to execute a step are iterative and should not be finalized until previous steps have been thoroughly addressed. The steps are as follows:

Step 1: Engage stakeholders.
Step 2: Describe the program.
Step 3: Focus the evaluation design.
Step 4: Gather credible evidence.
Step 5: Justify conclusions.
Step 6: Ensure use and share lessons learned.

Adhering to these six steps will facilitate an understanding of a program's context (e.g., the program's history, setting, and organization) and will improve how most evaluations are conceived and conducted.

The second element of the framework is a set of 30 standards for assessing the quality of evaluation activities, organized into the following four groups:

Standard 1: utility,
Standard 2: feasibility,
Standard 3: propriety, and
Standard 4: accuracy.

These standards, adopted from the Joint Committee on Standards for Educational Evaluation (*12*),* answer the question, "Will this evaluation be effective?" and are recommended as criteria for judging the quality of program evaluation efforts in public health. The remainder of this report discusses each step, its subpoints, and the standards that govern effective program evaluation (Box 1).

Steps in Program Evaluation

Step 1: Engaging Stakeholders

The evaluation cycle begins by engaging stakeholders (i.e., the persons or organizations having an investment in what will be learned from an evaluation and what will be done with the knowledge). Public health work involves partnerships; therefore, any assessment of a public health program requires considering the value systems of the partners. Stakeholders must be engaged in the inquiry to ensure that their perspectives are understood. When stakeholders are not engaged, an evaluation might not address important elements of a program's objectives, operations, and outcomes. Therefore, evaluation findings might be ignored, criticized, or resisted because the evaluation did not address the stakeholders' concerns or values (*12*). After becoming involved, stakeholders help to execute the other steps. Identifying and engaging the following three principal groups of stakeholders are critical:

*The program evaluation standards are an approved standard by the American National Standards Institute (ANSI) and have been endorsed by the American Evaluation Association and 14 other professional organizations (ANSI Standard No. JSEE-PR 1994, Approved March 15, 1994).

- those involved in program operations (e.g., sponsors, collaborators, coalition partners, funding officials, administrators, managers, and staff);
- those served or affected by the program (e.g., clients, family members, neighborhood organizations, academic institutions, elected officials, advocacy groups, professional associations, skeptics, opponents, and staff of related or competing organizations); and
- primary users of the evaluation.

Those Involved in Program Operations. Persons or organizations involved in program operations have a stake in how evaluation activities are conducted because the program might be altered as a result of what is learned. Although staff, funding officials, and partners work together on a program, they are not necessarily a single interest group. Subgroups might hold different perspectives and follow alternative agendas; furthermore, because these stakeholders have a professional role in the

BOX 1. Steps in evaluation practice and standards for effective evaluation

Steps in Evaluation Practice

- **Engage stakeholders**
  Those persons involved in or affected by the program and primary users of the evaluation.
- **Describe the program**
  Need, expected effects, activities, resources, stage, context, logic model.
- **Focus the evaluation design**
  Purpose, users, uses, questions, methods, agreements.
- **Gather credible evidence**
  Indicators, sources, quality, quantity, logistics.
- **Justify conclusions**
  Standards, analysis/synthesis, interpretation, judgment, recommendations.
- **Ensure use and share lessons learned**
  Design, preparation, feedback, follow-up, dissemination.

Standards for Effective Evaluation

- **Utility**
  Serve the information needs of intended users.
- **Feasibility**
  Be realistic, prudent, diplomatic, and frugal.
- **Propriety**
  Behave legally, ethically, and with regard for the welfare of those involved and those affected.
- **Accuracy**
  Reveal and convey technically accurate information.

program, they might perceive program evaluation as an effort to judge them personally. Program evaluation is related to but must be distinguished from personnel evaluation, which operates under different standards (*13*).

Those Served or Affected by the Program. Persons or organizations affected by the program, either directly (e.g., by receiving services) or indirectly (e.g., by benefitting from enhanced community assets), should be identified and engaged in the evaluation to the extent possible. Although engaging supporters of a program is natural, individuals who are openly skeptical or antagonistic toward the program also might be important stakeholders to engage. Opposition to a program might stem from differing values regarding what change is needed or how to achieve it. Opening an evaluation to opposing perspectives and enlisting the help of program opponents in the inquiry might be prudent because these efforts can strengthen the evaluation's credibility.

Primary Users of the Evaluation. Primary users of the evaluation are the specific persons who are in a position to do or decide something regarding the program. In practice, primary users will be a subset of all stakeholders identified. A successful evaluation will designate primary users early in its development and maintain frequent interaction with them so that the evaluation addresses their values and satisfies their unique information needs (*7*).

The scope and level of stakeholder involvement will vary for each program evaluation. Various activities reflect the requirement to engage stakeholders (Box 2) (*14*). For example, stakeholders can be directly involved in designing and conducting the evaluation. Also, they can be kept informed regarding progress of the evaluation through periodic meetings, reports, and other means of communication. Sharing power and resolving conflicts helps avoid overemphasis of values held by any specific stakeholder (*15*). Occasionally, stakeholders might be inclined to use their involvement in an evaluation to sabotage, distort, or discredit the program. Trust among stakeholders is essential; therefore, caution is required for preventing misuse of the evaluation process.

Step 2: Describing the Program

Program descriptions convey the mission and objectives of the program being evaluated. Descriptions should be sufficiently detailed to ensure understanding of program goals and strategies. The description should discuss the program's capacity to effect change, its stage of development, and how it fits into the larger organization and community. Program descriptions set the frame of reference for all subsequent decisions in an evaluation. The description enables comparisons with similar programs and facilitates attempts to connect program components to their effects (*12*). Moreover, stakeholders might have differing ideas regarding program goals and purposes. Evaluations done without agreement on the program definition are likely to be of limited use. Sometimes, negotiating with stakeholders to formulate a clear and logical description will bring benefits before data are available to evaluate program effectiveness (*7*). Aspects to include in a program description are need, expected effects, activities, resources, stage of development, context, and logic model.

Need. A statement of need describes the problem or opportunity that the program addresses and implies how the program will respond. Important features for describing a program's need include a) the nature and magnitude of the problem or

opportunity, b) which populations are affected, c) whether the need is changing, and d) in what manner the need is changing.

Expected Effects. Descriptions of expectations convey what the program must accomplish to be considered successful (i.e., program effects). For most programs, the effects unfold over time; therefore, the descriptions of expectations should be organized by time, ranging from specific (i.e., immediate) to broad (i.e., long-term) consequences. A program's mission, goals, and objectives all represent varying levels of specificity regarding a program's expectations. Also, forethought should be given to anticipate potential unintended consequences of the program.

Activities. Describing program activities (i.e., what the program does to effect change) permits specific steps, strategies, or actions to be arrayed in logical sequence. This demonstrates how each program activity relates to another and clarifies the program's hypothesized mechanism or theory of change (*16,17*). Also, program activity descriptions should distinguish the activities that are the direct responsibility of the program from those that are conducted by related programs or partners (*18*). External factors that might affect the program's success (e.g., secular trends in the community) should also be noted.

Resources. Resources include the time, talent, technology, equipment, information, money, and other assets available to conduct program activities. Program resource descriptions should convey the amount and intensity of program services and highlight situations where a mismatch exists between desired activities and resources available to execute those activities. In addition, economic evaluations require an understanding of all direct and indirect program inputs and costs (*19–21*).

BOX 2. Engaging stakeholders

Definition	Fostering input, participation, and power-sharing among those persons who have an investment in the conduct of the evalation and the findings; it is especially important to engage primary users of the evaluation.
Role	Helps increase chances that the evaluation will be useful; can improve the evaluation's credibility, clarify roles and responsibilities, enhance cultural competence, help protect human subjects, and avoid real or perceived conflicts of interest.

Example Activities

- Consulting insiders (e.g., leaders, staff, clients, and program funding sources) and outsiders (e.g., skeptics);
- Taking special effort to promote the inclusion of less powerful groups or individuals;
- Coordinating stakeholder input throughout the process of evaluation design, operation, and use; and
- Avoiding excessive stakeholder identification, which might prevent progress of the evaluation.

Adapted from Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

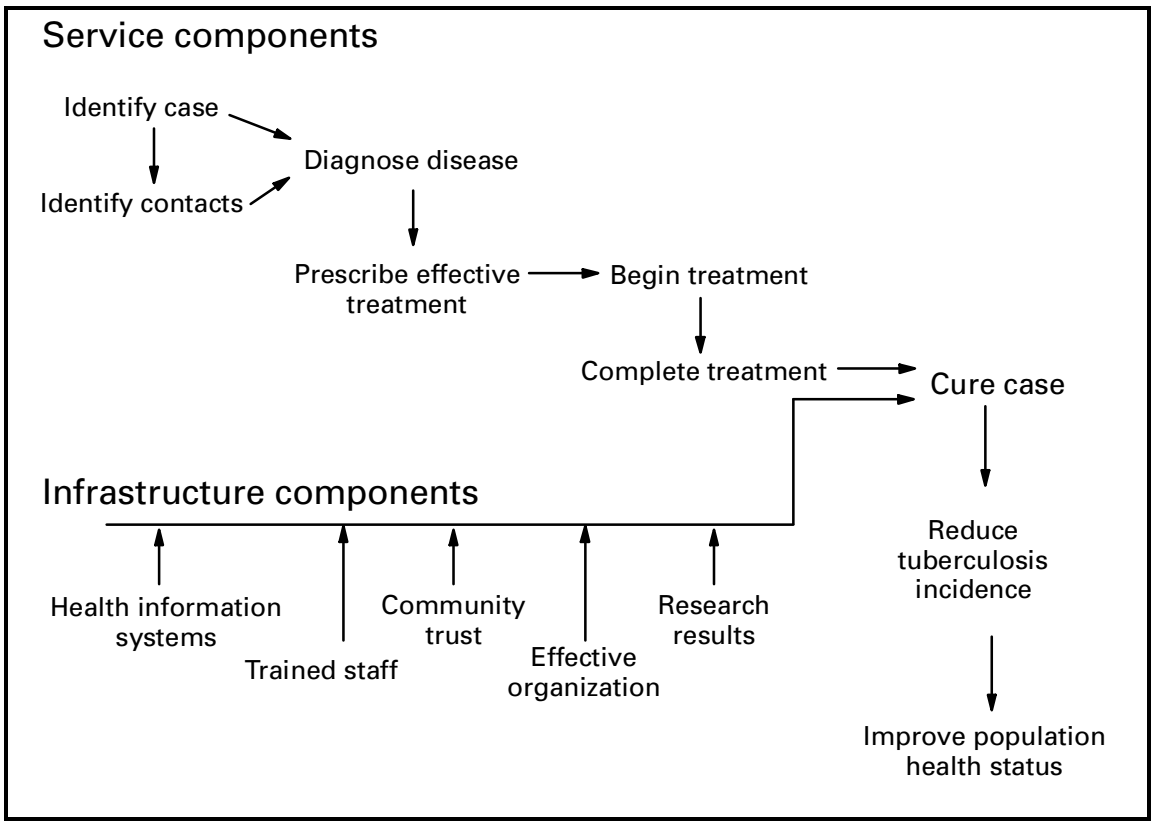
Stage of Development. Public health programs mature and change over time; therefore, a program's stage of development reflects its maturity. Programs that have recently received initial authorization and funding will differ from those that have been operating continuously for a decade. The changing maturity of program practice should be considered during the evaluation process (*22*). A minimum of three stages of development must be recognized: planning, implementation, and effects. During planning, program activities are untested, and the goal of evaluation is to refine plans. During implementation, program activities are being field-tested and modified; the goal of evaluation is to characterize real, as opposed to ideal, program activities and to improve operations, perhaps by revising plans. During the last stage, enough time has passed for the program's effects to emerge; the goal of evaluation is to identify and account for both intended and unintended effects.

Context. Descriptions of the program's context should include the setting and environmental influences (e.g., history, geography, politics, social and economic conditions, and efforts of related or competing organizations) within which the program operates (*6*). Understanding these environmental influences is required to design a context-sensitive evaluation and aid users in interpreting findings accurately and assessing the generalizability of the findings.

Logic Model. A logic model describes the sequence of events for bringing about change by synthesizing the main program elements into a picture of how the program is supposed to work (*23–35*). Often, this model is displayed in a flow chart, map, or table to portray the sequence of steps leading to program results (Figure 2). One of the virtues of a logic model is its ability to summarize the program's overall mechanism of change by linking processes (e.g., laboratory diagnosis of disease) to eventual effects (e.g., reduced tuberculosis incidence). The logic model can also display the infrastructure needed to support program operations. Elements that are connected within a logic model might vary but generally include inputs (e.g., trained staff), activities (e.g., identification of cases), outputs (e.g., persons completing treatment), and results ranging from immediate (e.g., curing affected persons) to intermediate (e.g., reduction in tuberculosis rate) to long-term effects (e.g., improvement of population health status). Creating a logic model allows stakeholders to clarify the program's strategies; therefore, the logic model improves and focuses program direction. It also reveals assumptions concerning conditions for program effectiveness and provides a frame of reference for one or more evaluations of the program. A detailed logic model can also strengthen claims of causality and be a basis for estimating the program's effect on endpoints that are not directly measured but are linked in a causal chain supported by prior research (*35*). Families of logic models can be created to display a program at different levels of detail, from different perspectives, or for different audiences.

Program descriptions will vary for each evaluation, and various activities reflect the requirement to describe the program (e.g., using multiple sources of information to construct a well-rounded description) (Box 3). The accuracy of a program description can be confirmed by consulting with diverse stakeholders, and reported descriptions of program practice can be checked against direct observation of activities in the field. A narrow program description can be improved by addressing such factors as staff turnover, inadequate resources, political pressures, or strong community participation that might affect program performance.

FIGURE 2. Logic model for a tuberculosis control program



Step 3: Focusing the Evaluation Design

The evaluation must be focused to assess the issues of greatest concern to stakeholders while using time and resources as efficiently as possible (*7,36,37*). Not all design options are equally well-suited to meeting the information needs of stakeholders. After data collection begins, changing procedures might be difficult or impossible, even if better methods become obvious. A thorough plan anticipates intended uses and creates an evaluation strategy with the greatest chance of being useful, feasible, ethical, and accurate. Among the items to consider when focusing an evaluation are purpose, users, uses, questions, methods, and agreements.

Purpose. Articulating an evaluation's purpose (i.e., intent) will prevent premature decision-making regarding how the evaluation should be conducted. Characteristics of the program, particularly its stage of development and context, will influence the evaluation's purpose. Public health evaluations have four general purposes. (Box 4). The first is to gain insight, which happens, for example, when assessing the feasibility of an innovative approach to practice. Knowledge from such an evaluation provides information concerning the practicality of a new approach, which can be used to design a program that will be tested for its effectiveness. For a developing program, information from prior evaluations can provide the necessary insight to clarify how its activities should be designed to bring about expected changes.

A second purpose for program evaluation is to change practice, which is appropriate in the implementation stage when an established program seeks to describe what it has done and to what extent. Such information can be used to better describe program processes, to improve how the program operates, and to fine-tune the overall program strategy. Evaluations done for this purpose include efforts to improve the quality, effectiveness, or efficiency of program activities.

A third purpose for evaluation is to assess effects. Evaluations done for this purpose examine the relationship between program activities and observed consequences. This type of evaluation is appropriate for mature programs that can define what interventions were delivered to what proportion of the target population. Knowing where to find potential effects can ensure that significant consequences are not overlooked. One set of effects might arise from a direct cause-and-effect relationship to the program. Where these exist, evidence can be found to attribute the effects exclusively to the program. In addition, effects might arise from a causal process involving issues of contribution as well as attribution. For example, if a program's activities are aligned with those of other programs operating in the same setting, certain effects (e.g., the creation of new laws or policies) cannot be attributed solely to one program or another. In such situations, the goal for evaluation is to gather credible

BOX 3. Describing the program

Definition Scrutinizing the features of the program being evaluated, including its purpose and place in a larger context. Description includes information regarding the way the program was intended to function and the way that it actually was implemented. Also includes features of the program's context that are likely to influence conclusions regarding the program.

Role Improves evaluation's fairness and accuracy; permits a balanced assessment of strengths and weaknesses and helps stakeholders understand how program features fit together and relate to a larger context.

Example Activities

- Characterizing the need (or set of needs) addressed by the program;
- Listing specific expectations as goals, objectives, and criteria for success;
- Clarifying why program activities are believed to lead to expected changes;
- Drawing an explicit logic model to illustrate relationships between program elements and expected changes;
- Assessing the program's maturity or stage of development;
- Analyzing the context within which the program operates;
- Considering how the program is linked to other ongoing efforts; and
- Avoiding creation of an overly precise description for a program that is under development.

Adapted from Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

evidence that describes each program's contribution in the combined change effort. Establishing accountability for program results is predicated on an ability to conduct evaluations that assess both of these kinds of effects.

A fourth purpose, which applies at any stage of program development, involves using the process of evaluation inquiry to affect those who participate in the inquiry. The logic and systematic reflection required of stakeholders who participate in an evaluation can be a catalyst for self-directed change. An evaluation can be initiated

BOX 4. Selected uses for evaluation in public health practice by category of purpose

Gain insight

- Assess needs, desires, and assets of community members.
- Identify barriers and facilitators to service use.
- Learn how to describe and measure program activities and effects.

Change practice

- Refine plans for introducing a new service.
- Characterize the extent to which intervention plans were implemented.
- Improve the content of educational materials.
- Enhance the program's cultural competence.
- Verify that participants' rights are protected.
- Set priorities for staff training.
- Make midcourse adjustments to improve patient/client flow.
- Improve the clarity of health communication messages.
- Determine if customer satisfaction rates can be improved.
- Mobilize community support for the program.

Assess effects

- Assess skills development by program participants.
- Compare changes in provider behavior over time.
- Compare costs with benefits.
- Find out which participants do well in the program.
- Decide where to allocate new resources.
- Document the level of success in accomplishing objectives.
- Demonstrate that accountability requirements are fulfilled.
- Aggregate information from several evaluations to estimate outcome effects for similar kinds of programs.
- Gather success stories.

Affect participants

- Reinforce intervention messages.
- Stimulate dialogue and raise awareness regarding health issues.
- Broaden consensus among coalition members regarding program goals.
- Teach evaluation skills to staff and other stakeholders.
- Support organizational change and development.

with the intent of generating a positive influence on stakeholders. Such influences might be to supplement the program intervention (e.g., using a follow-up questionnaire to reinforce program messages); empower program participants (e.g., increasing a client's sense of control over program direction); promote staff development (e.g., teaching staff how to collect, analyze, and interpret evidence); contribute to organizational growth (e.g., clarifying how the program relates to the organization's mission); or facilitate social transformation (e.g., advancing a community's struggle for self-determination) (*7,38–42*).

Users. Users are the specific persons that will receive evaluation findings. Because intended users directly experience the consequences of inevitable design trade-offs, they should participate in choosing the evaluation focus (*7*). User involvement is required for clarifying intended uses, prioritizing questions and methods, and preventing the evaluation from becoming misguided or irrelevant.

Uses. Uses are the specific ways in which information generated from the evaluation will be applied. Several uses exist for program evaluation (Box 4). Stating uses in vague terms that appeal to many stakeholders increases the chances the evaluation will not fully address anyone's needs. Uses should be planned and prioritized with input from stakeholders and with regard for the program's stage of development and current context. All uses must be linked to one or more specific users.

Questions. Questions establish boundaries for the evaluation by stating what aspects of the program will be addressed (*5–7*). Creating evaluation questions encourages stakeholders to reveal what they believe the evaluation should answer. Negotiating and prioritizing questions among stakeholders further refines a viable focus. The question-development phase also might expose differing stakeholder opinions regarding the best unit of analysis. Certain stakeholders might want to study how programs operate together as a system of interventions to effect change within a community. Other stakeholders might have questions concerning the performance of a single program or a local project within a program. Still others might want to concentrate on specific subcomponents or processes of a project. Clear decisions regarding the questions and corresponding units of analysis are needed in subsequent steps of the evaluation to guide method selection and evidence gathering.

Methods. The methods for an evaluation are drawn from scientific research options, particularly those developed in the social, behavioral, and health sciences (*5–7,43–48*). A classification of design types includes experimental, quasi-experimental, and observational designs (*43,48*). No design is better than another under all circumstances. Evaluation methods should be selected to provide the appropriate information to address stakeholders' questions (i.e., methods should be matched to the primary users, uses, and questions). Experimental designs use random assignment to compare the effect of an intervention with otherwise equivalent groups (*49*). Quasi-experimental methods compare nonequivalent groups (e.g., program participants versus those on a waiting list) or use multiple waves of data to set up a comparison (e.g., interrupted time series) (*50,51*). Observational methods use comparisons within a group to explain unique features of its members (e.g., comparative case studies or cross-sectional surveys) (*45,52–54*). The choice of design has implications for what will count as evidence, how that evidence will be gathered, and what kind of claims can be made (including the internal and external validity of conclusions) (*55*). Also, methodologic decisions clarify how the evaluation will operate (e.g., to what extent

program participants will be involved; how information sources will be selected; what data collection instruments will be used; who will collect the data; what data management systems will be needed; and what are the appropriate methods of analysis, synthesis, interpretation, and presentation). Because each method option has its own bias and limitations, evaluations that mix methods are generally more effective (*44,56–58*). During the course of an evaluation, methods might need to be revised or modified. Also, circumstances that make a particular approach credible and useful can change. For example, the evaluation's intended use can shift from improving a program's current activities to determining whether to expand program services to a new population group. Thus, changing conditions might require alteration or iterative redesign of methods to keep the evaluation on track (*22*).

Agreements. Agreements summarize the procedures and clarify roles and responsibilities among those who will execute the evaluation plan (*6,12*). Agreements describe how the evaluation plan will be implemented by using available resources (e.g., money, personnel, time, and information) (*36,37*). Agreements also state what safeguards are in place to protect human subjects and, where appropriate, what ethical (e.g., institutional review board) or administrative (e.g., paperwork reduction) approvals have been obtained (*59,60*). Elements of an agreement include statements concerning the intended purpose, users, uses, questions, and methods, as well as a summary of the deliverables, time line, and budget. The agreement can include all engaged stakeholders but, at a minimum, it must involve the primary users, any providers of financial or in-kind resources, and those persons who will conduct the evaluation and facilitate its use and dissemination. The formality of an agreement might vary depending on existing stakeholder relationships. An agreement might be a legal contract, a detailed protocol, or a memorandum of understanding. Creating an explicit agreement verifies the mutual understanding needed for a successful evaluation. It also provides a basis for modifying or renegotiating procedures if necessary.

Various activities reflect the requirement to focus the evaluation design (Box 5). Both supporters and skeptics of the program could be consulted to ensure that the proposed evaluation questions are politically viable (i.e., responsive to the varied positions of interest groups). A menu of potential evaluation uses appropriate for the program's stage of development and context could be circulated among stakeholders to determine which is most compelling. Interviews could be held with specific intended users to better understand their information needs and time line for action. Resource requirements could be reduced when users are willing to employ more timely but less precise evaluation methods.

Step 4: Gathering Credible Evidence

An evaluation should strive to collect information that will convey a well-rounded picture of the program so that the information is seen as credible by the evaluation's primary users. Information (i.e., evidence) should be perceived by stakeholders as believable and relevant for answering their questions. Such decisions depend on the evaluation questions being posed and the motives for asking them. For certain questions, a stakeholder's standard for credibility might require having the results of a controlled experiment; whereas for another question, a set of systematic observations (e.g., interactions between an outreach worker and community residents) would be the most credible. Consulting specialists in evaluation methodology might be

necessary in situations where concern for data quality is high or where serious consequences exist associated with making errors of inference (i.e., concluding that program effects exist when none do, concluding that no program effects exist when in fact they do, or attributing effects to a program that has not been adequately implemented) (*61,62*).

Having credible evidence strengthens evaluation judgments and the recommendations that follow from them. Although all types of data have limitations, an evaluation's overall credibility can be improved by using multiple procedures for gathering, analyzing, and interpreting data. Encouraging participation by stakeholders can also enhance perceived credibility. When stakeholders are involved in defining and gathering data that they find credible, they will be more likely to accept the evaluation's conclusions and to act on its recommendations (*7,38*). Aspects of evidence gathering

BOX 5. Focusing the evaluation design

Definition	Planning in advance where the evaluation is headed and what steps will be taken; process is iterative (i.e., it continues until a focused approach is found to answer evaluation questions with methods that stakeholders agree will be useful, feasible, ethical, and accurate); evaluation questions and methods might be adjusted to achieve an optimal match that facilitates use by primary users.
Role	Provides investment in quality; increases the chances that the evaluation will succeed by identifying procedures that are practical, politically viable, and cost-effective; failure to plan thoroughly can be self-defeating, leading to an evaluation that might become impractical or useless; when stakeholders agree on a design focus, it is used throughout the evaluation process to keep the project on track.

Example Activities

- Meeting with stakeholders to clarify the intent or purpose of the evaluation;
- Learning which persons are in a position to actually use the findings, then orienting the plan to meet their needs;
- Understanding how the evaluation results are to be used;
- Writing explicit evaluation questions to be answered;
- Describing practical methods for sampling, data collection, data analysis, interpretation, and judgment;
- Preparing a written protocol or agreement that summarizes the evaluation procedures, with clear roles and responsibilities for all stakeholders; and
- Revising parts or all of the evaluation plan when critical circumstances change.

Adapted from a) Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and b) U.S. General Accounting Office. Designing evaluations. Washington, DC: U.S. General Accounting Office, 1991; publication no. GAO/PEMD-10.1.4.

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

MMWR™
MORBIDITY AND MORTALITY WEEKLY REPORT

September 17, 1999 / Vol. 48 / No. RR-11

Recommendations and Reports

Continuing Education Activity Sponsored by CDC

Framework for Program Evaluation in Public Health

GOALS and OBJECTIVES

This *MMWR* provides recommendations regarding the conduct of public health program evaluation. These recommendations were developed by CDC staff and expert consultants. This report is intended to guide public health professionals in their use of program evaluation. Upon completion of this educational activity, the reader should be able to a) identify concepts that are essential for program evaluation; b) describe the purpose and features of the framework for program evaluation; c) identify and organize steps in program evaluation practice, as well as concepts that comprise each step; d) identify, organize, and know when to apply the standards for effective program evaluation; and e) discuss misconceptions regarding the purposes and methods of program evaluation.

ACCREDITATION

Continuing Medical Education (CME) Credit: This activity has been planned and implemented in accordance with the Essentials and Standards of the Accreditation Council for Continuing Medical Education (ACCME) through CDC. CDC is accredited by the ACCME to provide continuing medical education for physicians. CDC awards 2.0 hours of category 1 credit toward the AMA Physician's Recognition Award for this activity. Each physician should claim only those hours he/she actually spent in the educational activity.

Continuing Education Unit (CEU) Credit: CDC awards 0.2 hour of CEUs. This activity has been structured following the International Association for Continuing Education and Training (IACET) Criteria and Guidelines and therefore is awarding CEUs. The CEU is a nationally recognized unit designed to provide a record of an individual's continuing education accomplishments.

Continuing Nursing Education (CNE) Credit: This activity for 2.6 contact hours is provided by CDC, which is accredited as a provider of continuing nursing education by the American Nurses Credentialing Center's (ANCC) Commission on Accreditation.

EXPIRATION — September 17, 2000

The response form must be completed and returned electronically, by fax, or by mail, **postmarked no later than 1 year from the publication date of this report**, for eligibility to receive continuing education credit.

INSTRUCTIONS

1. Read this *MMWR* (Vol. 48, RR-11), which contains the correct answers to the questions beginning on the next page.
2. Complete all registration information on the response form, including your name, mailing address, phone number, and e-mail address, if available.
3. Indicate whether you are registering for Continuing Medical Education (CME) credit, Continuing Education Unit (CEU) credit, or Continuing Nursing Education (CNE) credit.
4. Select your answers to the questions, and mark the corresponding letters on the response form. To receive continuing education credit, you must answer *all* of the questions. Questions with more than one correct answer will instruct you to "indicate all that are true."
5. Sign and date the response form.
6. Return the response form, or a photocopy of the form, no later than September 17, 2000, to CDC by one of the following methods:

   Internet: <http://www.cdc.gov/mmwr>
   Fax: 404-639-4198

   Mail: MMWR CE Credit
   Office of Scientific and Health Communications
   Epidemiology Program Office — MS C08
   Centers for Disease Control and Prevention
   1600 Clifton Road, N.E.
   Atlanta, GA 30333

   If you answer all of the questions, you will receive an award letter for 2.0 hours of CME credit, 0.2 hour of CEU credit, or 2.6 hours of CNE credit within 90 days. No fees are charged for participating in this continuing education activity.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICES

To receive continuing education credit, please answer all of the following questions. For each question, indicate the one best answer.

1. Formal evaluation procedures become important when . . .

A. developing staffing plans for a program.

B. making small changes in program functions.

C. the stakes of potential decisions or program changes increase.

D. altering outreach procedures to increase program participation.

2. Which of the following is NOT generally addressed by value questions?

A. Worth.

B. Merit.

C. Justice.

D. Significance.

3. The framework for program evaluation was designed to . . .

A. standardize the way public health professionals conduct program evaluations.

B. guide public health professionals in their use of program evaluation.

C. improve the accuracy of program evaluation findings.

D. clarify new responsibilities for public health professionals.

4. Which of the following is NOT one of the steps of program evaluation?

A. Gathering credible evidence.

B. Comparing with baseline data.

C. Engaging stakeholders.

D. Justifying conclusions.

5. If lessons learned from an evaluation are not used, the evaluation should be considered . . .

A. ineffective.

B. unethical.

C. censored.

D. political.

6. Which step in the framework for program evaluation involves clarifying the logical sequence that links program activities with their intended effects?

A. Justifying conclusions.

B. Focusing the evaluation design.

C. Gathering credible evidence.

D. Describing the program.

7. Using hypothetical data to rehearse how evaluation findings could be used is an example of which concept in the framework for program evaluation?

A. Expected effects.

B. Preparation.

C. Dissemination.

D. Interpretation.

8. The standards for effective evaluation state that the evaluation should be . . .

A. systematic, fair, helpful, and cost-effective.

B. useful, feasible, ethical, and accurate.

C. reasonable, unbiased, confidential, and well-planned.

D. comprehensive, reliable, worthwhile, and unobtrusive.

9. The standards for effective program evaluation should be applied . . .

A. at the end of an evaluation project.

B. at the beginning of an evaluation project.

C. while the evaluation is being planned and throughout its implementation.

D. when a new group of stakeholders becomes engaged in the evaluation.

10. Which of the following is NOT inherent within the practical approach encouraged by the framework for program evaluation?

A. A collaborative, team approach.

B. Beginning evaluation as early as possible in the life of a program.

C. Designing evaluations to achieve intended uses by primary users.

D. Using precise methods of analysis to quantify program impact.

11. Indicate your work setting.

A. State/local health department.
B. Other public health setting.
C. Hospital clinic/private practice.
D. Managed care organization.
E. Academic institution.
F. Other.

12. Which best describes your professional activities?

A. Patient care — emergency/urgent care department.
B. Patient care — inpatient.
C. Patient care — primary-care clinic.
D. Laboratory/pharmacy.
E. Administration.
F. Public health.

13. I plan to use these guidelines as the basis for . . . (Indicate all that apply.)

A. health education materials.
B. insurance reimbursement policies.
C. local practice guidelines.
D. public policy.
E. other.

14. How much time did you spend reading this report and completing the exam?

A. 1–1½ hours.
B. More than 1½ hours but fewer than 2 hours.
C. 2–2½ hours.
D. More than 2½ hours.

15. After reading this report, I am confident I can identify concepts that are essential for program evaluation.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

16. After reading this report, I am confident I can describe the purpose and features of the framework for program evaluation.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

17. After reading this report, I am confident I can identify and organize steps in program evaluation practice, as well as concepts that comprise each step.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

18. After reading this report, I am confident I can identify, organize, and know when to apply the standards for effective program evaluation.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

19. After reading this report, I am confident I can discuss misconceptions regarding the purposes and methods of program evaluation.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

20. The text boxes and figures are useful.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

21. Overall, the presentation of the report enhanced my ability to understand the material.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

22. These recommendations will affect how I conduct or participate in program evaluations.

A. Strongly agree.

B. Agree.

C. Neither agree nor disagree.

D. Disagree.

E. Strongly disagree.

Correct answers for questions 1-10

1. C; 2. C; 3. B; 4. B; 5. A; 6. D; 7. B; 8. B; 9. C; 10. D.

MMWR Response Form for Continuing Education Credit
September 17, 1999 / Vol. 48 / No. RR-11

Framework for Program Evaluation in Public Health

Fill in the appropriate block(s) to indicate your answer(s).
To receive continuing education credit, you must answer all of the questions.

1.	[ ] A	[ ] B	[ ] C	[ ] D		
2.	[ ] A	[ ] B	[ ] C	[ ] D		
3.	[ ] A	[ ] B	[ ] C	[ ] D		
4.	[ ] A	[ ] B	[ ] C	[ ] D		
5.	[ ] A	[ ] B	[ ] C	[ ] D		
6.	[ ] A	[ ] B	[ ] C	[ ] D		
7.	[ ] A	[ ] B	[ ] C	[ ] D		
8.	[ ] A	[ ] B	[ ] C	[ ] D		
9.	[ ] A	[ ] B	[ ] C	[ ] D		
10.	[ ] A	[ ] B	[ ] C	[ ] D		
11.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	[ ] F
12.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	[ ] F
13.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
14.	[ ] A	[ ] B	[ ] C	[ ] D		
15.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
16.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
17.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
18.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
19.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
20.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
21.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	
22.	[ ] A	[ ] B	[ ] C	[ ] D	[ ] E	

Detach or photocopy.

Please Print:

Name: _____

Address: _____

_____

Telephone No.: _____ E-mail: _____

Fax No.: _____

Check one box below:

- ☐ 2.0 hours of CME credit
- ☐ 0.2 hour of CEU credit
- ☐ 2.6 hours of CNE credit

I completed this exam on

_____
(Date)

that typically affect perceptions of credibility include indicators, sources, quality, quantity, and logistics.

Indicators. Indicators define the program attributes that pertain to the evaluation's focus and questions (*63–66*). Because indicators translate general concepts regarding the program, its context, and its expected effects into specific measures that can be interpreted, they provide a basis for collecting evidence that is valid and reliable for the evaluation's intended uses. Indicators address criteria that will be used to judge the program; therefore, indicators reflect aspects of the program that are meaningful for monitoring (*66–70*). Examples of indicators that can be defined and tracked include measures of program activities (e.g., the program's capacity to deliver services; the participation rate; levels of client satisfaction; the efficiency of resource use; and the amount of intervention exposure) and measures of program effects (e.g., changes in participant behavior, community norms, policies or practices, health status, quality of life, and the settings or environment around the program).

Defining too many indicators can detract from the evaluation's goals; however, multiple indicators are needed for tracking the implementation and effects of a program. One approach to developing multiple indicators is based on the program logic model (developed in the second step of the evaluation). The logic model can be used as a template to define a spectrum of indicators leading from program activities to expected effects (*23,29–35*). For each step in the model, qualitative/quantitative indicators could be developed to suit the concept in question, the information available, and the planned data uses. Relating indicators to the logic model allows the detection of small changes in performance faster than if a single outcome were the only measure used. Lines of responsibility and accountability are also clarified through this approach because the measures are aligned with each step of the program strategy. Further, this approach results in a set of broad-based measures that reveal how health outcomes are the consequence of intermediate effects of the program. Intangible factors (e.g., service quality, community capacity [*71*], or interorganizational relations) that also affect the program can be measured by systematically recording markers of what is said or done when the concept is expressed (*72,73*). During an evaluation, indicators might need to be modified or new ones adopted. Measuring program performance by tracking indicators is only part of an evaluation and must not be confused as a singular basis for decision-making. Well-documented problems result from using performance indicators as a substitute for completing the evaluation process and reaching fully justified conclusions (*66,67,74*). An indicator (e.g., a rising rate of disease) might be assumed to reflect a failing program when, in reality, the indicator might be influenced by changing conditions that are beyond the program's control.

Sources. Sources of evidence in an evaluation are the persons, documents, or observations that provide information for the inquiry (Box 6). More than one source might be used to gather evidence for each indicator to be measured. Selecting multiple sources provides an opportunity to include different perspectives regarding the program and thus enhances the evaluation's credibility. An inside perspective might be understood from internal documents and comments from staff or program managers, whereas clients, neutral observers, or those who do not support the program might provide a different, but equally relevant perspective. Mixing these and other perspectives provides a more comprehensive view of the program. The criteria used for selecting sources should be stated clearly so that users and other stakeholders can

interpret the evidence accurately and assess if it might be biased (*45,75–77*). In addition, some sources are narrative in form and others are numeric. The integration of qualitative and quantitative information can increase the chances that the evidence base will be balanced, thereby meeting the needs and expectations of diverse users (*43,45,56,57,78–80*). Finally, in certain cases, separate evaluations might be selected as sources for conducting a larger synthesis evaluation (*58,81,82*).

Quality. Quality refers to the appropriateness and integrity of information used in an evaluation. High-quality data are reliable, valid, and informative for their intended use. Well-defined indicators enable easier collection of quality data. Other factors affecting quality include instrument design, data-collection procedures, training of

BOX 6. Selected sources of evidence for an evaluation

Persons

- Clients, program participants, nonparticipants;
- Staff, program managers, administrators;
- General public;
- Key informants;
- Funding officials;
- Critics/skeptics;
- Staff of other agencies;
- Representatives of advocacy groups;
- Elected officials, legislators, policymakers; and
- Local and state health officials.

Documents

- Grant proposals, newsletters, press releases;
- Meeting minutes, administrative records, registration/enrollment forms;
- Publicity materials, quarterly reports;
- Publications, journal articles, posters;
- Previous evaluation reports;
- Asset and needs assessments;
- Surveillance summaries;
- Database records;
- Records held by funding officials or collaborators;
- Internet pages; and
- Graphs, maps, charts, photographs, videotapes.

Observations

- Meetings, special events/activities, job performance; and
- Service encounters.

Adapted from Taylor-Powell E, Rossing B, Geran J. Evaluating collaboratives: reaching the potential. Madison, WI: University of Wisconsin Cooperative Extension, 1998.

data collectors, source selection, coding, data management, and routine error checking. Obtaining quality data will entail trade-offs (e.g., breadth versus depth) that should be negotiated among stakeholders. Because all data have limitations, the intent of a practical evaluation is to strive for a level of quality that meets the stakeholders' threshold for credibility.

Quantity. Quantity refers to the amount of evidence gathered in an evaluation. The amount of information required should be estimated in advance, or where evolving processes are used, criteria should be set for deciding when to stop collecting data. Quantity affects the potential confidence level or precision of the evaluation's conclusions. It also partly determines whether the evaluation will have sufficient power to detect effects (*83*). All evidence collected should have a clear, anticipated use. Correspondingly, only a minimal burden should be placed on respondents for providing information.

Logistics. Logistics encompass the methods, timing, and physical infrastructure for gathering and handling evidence. Each technique selected for gathering evidence (Box 7) must be suited to the source(s), analysis plan, and strategy for communicating findings. Persons and organizations also have cultural preferences that dictate acceptable ways of asking questions and collecting information, including who would be perceived as an appropriate person to ask the questions. For example, some participants might be willing to discuss their health behavior with a stranger, whereas others are more at ease with someone they know. The procedures for gathering evidence in an evaluation (Box 8) must be aligned with the cultural conditions in each setting of the project and scrutinized to ensure that the privacy and confidentiality of the information and sources are protected (*59,60,84*).

Step 5: Justifying Conclusions

The evaluation conclusions are justified when they are linked to the evidence gathered and judged against agreed-upon values or standards set by the stakeholders. Stakeholders must agree that conclusions are justified before they will use the evaluation results with confidence. Justifying conclusions on the basis of evidence includes standards, analysis and synthesis, interpretation, judgment, and recommendations.

Standards. Standards reflect the values held by stakeholders, and those values provide the basis for forming judgments concerning program performance. Using explicit standards distinguishes evaluation from other approaches to strategic management in which priorities are set without reference to explicit values. In practice, when stakeholders articulate and negotiate their values, these become the standards for judging whether a given program's performance will, for example, be considered successful, adequate, or unsuccessful. An array of value systems might serve as sources of norm-referenced or criterion-referenced standards (Box 9). When operationalized, these standards establish a comparison by which the program can be judged (*3,7,12*).

Analysis and Synthesis. Analysis and synthesis of an evaluation's findings might detect patterns in evidence, either by isolating important findings (analysis) or by combining sources of information to reach a larger understanding (synthesis). Mixed method evaluations require the separate analysis of each evidence element and a synthesis of all sources for examining patterns of agreement, convergence, or complexity. Deciphering facts from a body of evidence involves deciding how to organize, classify, interrelate, compare, and display information (*7,85–87*). These decisions are

BOX 7. Selected techniques for gathering evidence

- Written survey (e.g. handout, telephone, fax, mail, e-mail, or Internet);
- Personal interview (e.g. individual or group; structured, semistructured, or conversational);
- Observation;
- Document analysis;
- Case study;
- Group assessment (e.g. brainstorming or nominal group [i.e., a structured group process conducted to elicit and rank priorities, set goals, or identify problems]);
- Role play, dramatization;
- Expert or peer review;
- Portfolio review;
- Testimonials;
- Semantic differentials, paired comparisons, similarity or dissimilarity tests;
- Hypothetical scenarios;
- Storytelling;
- Geographical mapping;
- Concept mapping;
- Pile sorting (i.e., a technique that allows respondents to freely categorize items, revealing how hey perceive the structure of a domain);
- Free-listing (i.e., a technique to elicit a complete list of all items in a cultural domain);
- Social network diagraming;
- Simulation, modeling;
- Debriefing sessions;
- Cost accounting;
- Photography, drawing, art, videography;
- Diaries or journals; and
- Logs, activity forms, registries.

Adapted from a) Taylor-Powell E, Rossing B, Geran J. Evaluating collaboratives: reaching the potential. Madison, WI: University of Wisconsin Cooperative Extension, 1998; b) Phillips JJ. Handbook of training evaluation and measurement methods. 3rd ed. Houston, TX: Gulf Publishing Company, 1997; c) Weller SC. Systematic data collection. Thousand Oaks, CA: Sage Publications, Inc. 1988; and d) Trochim WMK. Introduction to concept mapping for planning and evaluation. Available at <http://trochim.human.cornell.edu/research/epp1/epp1.htm>. Accessed July 1999.

BOX 8. Gathering credible evidence

Definition Compiling information that stakeholders perceive as trustworthy and 'relevant for answering their questions. Such evidence can be experimental or observational, qualitative or quantitative, or it can include a mixture of methods. Adequate data might be available and easily accessed, or it might need to be defined and new data collected. Whether a body of evidence is credible to stakeholders depends on such factors as how the questions were posed, sources of information, conditions of data collection, reliability of measurement, validity of interpretations, and quality control procedures.

Role Enhances the evaluation's utility and accuracy; guides the scope and selection of information and gives priority to the most defensible information sources; promotes the collection of valid, reliable, and systematic information that is the foundation of any effective evaluation.

Example Activities

- Choosing indicators that meaningfully address evaluation questions;
- Describing fully the attributes of information sources and the rationale for their selection;
- Establishing clear procedures and training staff to collect high-quality information;
- Monitoring periodically the quality of information obtained and taking practical steps to improve quality;
- Estimating in advance the amount of information required or establishing criteria for deciding when to stop collecting data in situations where an iterative or evolving process is used; and
- Safeguarding the confidentiality of information and information sources.

Adapted from Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

guided by the questions being asked, the types of data available, and by input from stakeholders and primary users.

Interpretation. Interpretation is the effort of figuring out what the findings mean and is part of the overall effort to understand the evidence gathered in an evaluation (*88*). Uncovering facts regarding a program's performance is not sufficient to draw evaluative conclusions. Evaluation evidence must be interpreted to determine the practical significance of what has been learned. Interpretations draw on information and perspectives that stakeholders bring to the evaluation inquiry and can be strengthened through active participation or interaction.

Judgments. Judgments are statements concerning the merit, worth, or significance of the program. They are formed by comparing the findings and interpretations

regarding the program against one or more selected standards. Because multiple standards can be applied to a given program, stakeholders might reach different or even conflicting judgments. For example, a program that increases its outreach by 10% from the previous year might be judged positively by program managers who are

BOX 9. Selected sources of standards for judging program performance

- Needs of participants;
- Community values, expectations, norms;
- Degree of participation;
- Program objectives;
- Program protocols and procedures;
- Expected performance, forecasts, estimates;
- Feasibility;
- Sustainability;
- Absence of harms;
- Targets or fixed criteria of performance;
- Change in performance over time;
- Performance by previous or similar programs;
- Performance by a control or comparison group;
- Resource efficiency;
- Professional standards;
- Mandates, policies, statutes, regulations, laws;
- Judgments by reference groups (e.g., participants, staff, experts, and funding officials);
- Institutional goals;
- Political ideology;
- Social equity;
- Political will; and
- Human rights.

Adapted from a) Patton MQ. Utilization-focused evaluation: the new century text. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997; b) Scriven M. Minimalist theory of evaluation: the least theory that practice requires. American Journal of Evaluation 1998;19(1):57–70; c) McKenzie JF. Planning, implementing, and evaluating health promotion programs: a primer. New York, NY: Macmillan Publishing Company, 1993; d) Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and e) Gostin L, Mann JM. Towards the development of a human rights impact assessment for the formulation and evaluation of public health policies. Health and Human Rights 1994;1:59–80.

using the standard of improved performance over time. However, community members might feel that despite improvements, a minimum threshold of access to services has not been reached. Therefore, by using the standard of social equity, their judgment concerning program performance would be negative. Conflicting claims regarding a program's quality, value, or importance often indicate that stakeholders are using different standards for judgment. In the context of an evaluation, such disagreement can be a catalyst for clarifying relevant values and for negotiating the appropriate bases on which the program should be judged.

Recommendations. Recommendations are actions for consideration resulting from the evaluation. Forming recommendations is a distinct element of program evaluation that requires information beyond what is necessary to form judgments regarding program performance (*3*). Knowing that a program is able to reduce the risk of disease does not translate necessarily into a recommendation to continue the effort, particularly when competing priorities or other effective alternatives exist. Thus, recommendations for continuing, expanding, redesigning, or terminating a program are separate from judgments regarding a program's effectiveness. Making recommendations requires information concerning the context, particularly the organizational context, in which programmatic decisions will be made (*89*). Recommendations that lack sufficient evidence or those that are not aligned with stakeholders' values can undermine an evaluation's credibility. By contrast, an evaluation can be strengthened by recommendations that anticipate the political sensitivities of intended users and highlight areas that users can control or influence (*7*). Sharing draft recommendations, soliciting reactions from multiple stakeholders, and presenting options instead of directive advice increase the likelihood that recommendations will be relevant and well-received.

Various activities fulfill the requirement for justifying conclusions in an evaluation (Box 10). Conclusions could be strengthened by a) summarizing the plausible mechanisms of change; b) delineating the temporal sequence between activities and effects; c) searching for alternative explanations and showing why they are unsupported by the evidence; and d) showing that the effects can be repeated. When different but equally well-supported conclusions exist, each could be presented with a summary of its strengths and weaknesses. Creative techniques (e.g., the Delphi process*) could be used to establish consensus among stakeholders when assigning value judgments (*90*). Techniques for analyzing, synthesizing, and interpreting findings should be agreed on before data collection begins to ensure that all necessary evidence will be available.

Step 6: Ensuring Use and Sharing Lessons Learned

Lessons learned in the course of an evaluation do not automatically translate into informed decision-making and appropriate action. Deliberate effort is needed to ensure that the evaluation processes and findings are used and disseminated appropriately. Preparing for use involves strategic thinking and continued vigilance, both of

*Developed by the Rand Corporation, the Delphi process is an iterative method for arriving at a consensus concerning an issue or problem by circulating questions and responses to a panel of qualified reviewers whose identities are usually not revealed to one another. The questions and responses are progressively refined with each round until a viable option or solution is reached.

which begin in the earliest stages of stakeholder engagement and continue throughout the evaluation process. Five elements are critical for ensuring use of an evaluation, including design, preparation, feedback, follow-up, and dissemination.

Design. Design refers to how the evaluation's questions, methods, and overall processes are constructed. As discussed in the third step of this framework, the design should be organized from the start to achieve intended uses by primary users. Having a clear design that is focused on use helps persons who will conduct the evaluation to know precisely who will do what with the findings and who will benefit from being a part of the evaluation. Furthermore, the process of creating a clear design will highlight ways that stakeholders, through their contributions, can enhance the relevance, credibility, and overall utility of the evaluation.

Preparation. Preparation refers to the steps taken to rehearse eventual use of the evaluation findings. The ability to translate new knowledge into appropriate action is a skill that can be strengthened through practice. Building this skill can itself be a useful benefit of the evaluation (*38,39,91*). Rehearsing how potential findings

BOX 10. Justifying conclusions

Definition	Making claims regarding the program that are warranted on the basis of data that have been compared against pertinent and defensible ideas of merit, worth, or significance (i.e., against standards of values); conclusions are justified when they are linked to the evidence gathered and consistent with the agreed on values or standards of stakeholders.
Role	Reinforces conclusions central to the evaluation's utility and accuracy; involves values clarification, qualitative and quantitative data analysis and synthesis, systematic interpretation, and appropriate comparison against relevant standards for judgment.

Example Activities

- Using appropriate methods of analysis and synthesis to summarize findings;
- Interpreting the significance of results for deciding what the findings mean;
- Making judgments according to clearly stated values that classify a result (e.g., as positive or negative and high or low);
- Considering alternative ways to compare results (e.g., compared with program objectives, a comparison group, national norms, past performance, or needs);
- Generating alternative explanations for findings and indicating why these explanations should or should not be discounted;
- Recommending actions or decisions that are consistent with the conclusions; and
- Limiting conclusions to situations, time periods, persons, contexts, and purposes for which the findings are applicable.

Adapted from Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

(particularly negative findings) might affect decision-making will prepare stakeholders for eventually using the evidence (*92*). Primary users and other stakeholders could be given a set of hypothetical results and asked to explain what decisions or actions they would make on the basis of this new knowledge. If they indicate that the evidence presented is incomplete and that no action would be taken, this is a sign that the planned evaluation should be modified. Preparing for use also gives stakeholders time to explore positive and negative implications of potential results and time to identify options for program improvement.

Feedback. Feedback is the communication that occurs among all parties to the evaluation. Giving and receiving feedback creates an atmosphere of trust among stakeholders; it keeps an evaluation on track by letting those involved stay informed regarding how the evaluation is proceeding. Primary users and other stakeholders have a right to comment on decisions that might affect the likelihood of obtaining useful information. Stakeholder feedback is an integral part of evaluation, particularly for ensuring use. Obtaining feedback can be encouraged by holding periodic discussions during each step of the evaluation process and routinely sharing interim findings, provisional interpretations, and draft reports.

Follow-Up. Follow-up refers to the technical and emotional support that users need during the evaluation and after they receive evaluation findings. Because of the effort required, reaching justified conclusions in an evaluation can seem like an end in itself; however, active follow-up might be necessary to remind intended users of their planned use. Follow-up might also be required to prevent lessons learned from becoming lost or ignored in the process of making complex or politically sensitive decisions. To guard against such oversight, someone involved in the evaluation should serve as an advocate for the evaluation's findings during the decision-making phase. This type of advocacy increases appreciation of what was discovered and what actions are consistent with the findings.

Facilitating use of evaluation findings also carries with it the responsibility for preventing misuse (*7,12,74,93,94*). Evaluation results are always bound by the context in which the evaluation was conducted. However, certain stakeholders might be tempted to take results out of context or to use them for purposes other than those agreed on. For instance, inappropriately generalizing the results from a single case study to make decisions that affect all sites in a national program would constitute misuse of the case study evaluation. Similarly, stakeholders seeking to undermine a program might misuse results by overemphasizing negative findings without giving regard to the program's positive attributes. Active follow-up might help prevent these and other forms of misuse by ensuring that evidence is not misinterpreted and is not applied to questions other than those that were the central focus of the evaluation.

Dissemination. Dissemination is the process of communicating either the procedures or the lessons learned from an evaluation to relevant audiences in a timely, unbiased, and consistent fashion. Although documentation of the evaluation is needed, a formal evaluation report is not always the best or even a necessary product. Like other elements of the evaluation, the reporting strategy should be discussed in advance with intended users and other stakeholders. Such consultation ensures that the information needs of relevant audiences will be met. Planning effective communication also requires considering the timing, style, tone, message source, vehicle, and format of information products. Regardless of how communications are constructed,

the goal for dissemination is to achieve full disclosure and impartial reporting. A checklist of items to consider when developing evaluation reports includes tailoring the report content for the audience, explaining the focus of the evaluation and its limitations, and listing both the strengths and weaknesses of the evaluation (Box 11) (*6*).

Additional Uses. Additional uses for evaluation flow from the process of conducting the evaluation; these process uses have value and should be encouraged because they complement the uses of the evaluation findings (Box 12) (*7,38,93,94*). Those persons who participate in an evaluation can experience profound changes in thinking and behavior. In particular, when newcomers to evaluation begin to think as evaluators, fundamental shifts in perspective can occur. Evaluation prompts staff to clarify their understanding of program goals. This greater clarity allows staff to function cohesively as a team, focused on a common end (*95*). Immersion in the logic, reasoning, and values of evaluation can lead to lasting impacts (e.g., basing decisions on systematic judgments instead of on unfounded assumptions) (*7*). Additional process

BOX 11. Checklist for ensuring effective evaluation reports

- Provide interim and final reports to intended users in time for use.
- Tailor the report content, format, and style for the audience(s) by involving audience members.
- Include a summary.
- Summarize the description of the stakeholders and how they were engaged.
- Describe essential features of the program (e.g., including logic models).
- Explain the focus of the evaluation and its limitations.
- Include an adequate summary of the evaluation plan and procedures.
- Provide all necessary technical information (e.g., in appendices).
- Specify the standards and criteria for evaluative judgments.
- Explain the evaluative judgments and how they are supported by the evidence.
- List both strengths and weaknesses of the evaluation.
- Discuss recommendations for action with their advantages, disadvantages, and resource implications.
- Ensure protections for program clients and other stakeholders.
- Anticipate how people or organizations might be affected by the findings.
- Present minority opinions or rejoinders where necessary.
- Verify that the report is accurate and unbiased.
- Organize the report logically and include appropriate details.
- Remove technical jargon.
- Use examples, illustrations, graphics, and stories.

Adapted from Worthen BR, Sanders JR, Fitzpatrick JL. Program evaluation: alternative approaches and practical guidelines. 2nd ed. New York, NY: Logman, Inc. 1996.

uses for evaluation includes defining indicators to discover what matters to decision makers and making outcomes matter by changing the structural reinforcements connected with outcome attainment (e.g., by paying outcome dividends to programs that save money through their prevention efforts) (*96*). The benefits that arise from these and other process uses provide further rationale for initiating evaluation activities at the beginning of a program.

Standards for Effective Evaluation

Public health professionals will recognize that the basic steps of the framework for program evaluation are part of their routine work. In day-to-day public health practice, stakeholders are consulted; program goals are defined; guiding questions are stated; data are collected, analyzed, and interpreted; judgments are formed; and lessons are shared. Although informal evaluation occurs through routine practice, standards exist to assess whether a set of evaluative activities are well-designed and working to their potential. The Joint Committee on Standards for Educational

BOX 12. Ensuring use and sharing lessons learned

Definition Ensuring that a) stakeholders are aware of the evaluation procedures and findings; b) the findings are considered in decisions or actions that affect the program (i.e., findings use); and c) those who participated in the evaluation have had a beneficial experience (i.e., process use).

Role Ensures that evaluation achieves its primary purpose — being useful; however, several factors might influence the degree of use, including evaluator credibility, report clarity, report timeliness and dissemination, disclosure of findings, impartial reporting, and changes in the program or organization context.

Example Activities

- Designing the evaluation to achieve intended use by intended users;
- Preparing stakeholders for eventual use by rehearsing throughout the project how different kinds of conclusions would affect program operations;
- Providing continuous feedback to stakeholders regarding interim findings, provisional interpretations, and decisions to be made that might affect likelihood of use;
- Scheduling follow-up meetings with intended users to facilitate the transfer of evaluation conclusions into appropriate actions or decisions; and
- Disseminating both the procedures used and the lessons learned from the evaluation to stakeholders, using tailored communications strategies that meet their particular needs.

Adapted from a) Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994; and b) Patton MQ. Utilization-focused evaluation: the new century text. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997.

Evaluation has developed program evaluation standards for this purpose (*12*). These standards, designed to assess evaluations of educational programs, are also relevant for public health programs.

The program evaluation standards make conducting sound and fair evaluations practical. The standards provide practical guidelines to follow when having to decide among evaluation options. The standards help avoid creating an imbalanced evaluation (e.g., one that is accurate and feasible but not useful or one that would be useful and accurate but is infeasible). Furthermore, the standards can be applied while planning an evaluation and throughout its implementation. The Joint Committee is unequivocal in that, "the standards are guiding principles, not mechanical rules. . . . In the end, whether a given standard has been addressed adequately in a particular situation is a matter of judgment" (*12*).

In the Joint Committee's report, standards are grouped into the following four categories and include a total of 30 specific standards (Boxes 13–16). As described in the report, each category has an associated list of guidelines and common errors, illustrated with applied case examples:

- utility,
- feasibility,
- propriety, and
- accuracy.

Standard 1: Utility

Utility standards ensure that information needs of evaluation users are satisfied. Seven utility standards (Box 13) address such items as identifying those who will be impacted by the evaluation, the amount and type of information collected, the values used in interpreting evaluation findings, and the clarity and timeliness of evaluation reports.

Standard 2: Feasibility

Feasibility standards ensure that the evaluation is viable and pragmatic. The three feasibility standards (Box 14) emphasize that the evaluation should employ practical, nondisruptive procedures; that the differing political interests of those involved should be anticipated and acknowledged; and that the use of resources in conducting the evaluation should be prudent and produce valuable findings.

Standard 3: Propriety

Propriety standards ensure that the evaluation is ethical (i.e., conducted with regard for the rights and interests of those involved and effected). Eight propriety standards (Box 15) address such items as developing protocols and other agreements for guiding the evaluation; protecting the welfare of human subjects; weighing and disclosing findings in a complete and balanced fashion; and addressing any conflicts of interest in an open and fair manner.

BOX 13. Utility standards

The following utility standards ensure that an evaluation will serve the information needs of intended users:

A. **Stakeholder identification.** Persons involved in or affected by the evaluation should be identified so that their needs can be addressed.
B. **Evaluator credibility.** The persons conducting the evaluation should be trustworthy and competent in performing the evaluation for findings to achieve maximum credibility and acceptance.
C. **Information scope and selection.** Information collected should address pertinent questions regarding the program and be responsive to the needs and interests of clients and other specified stakeholders.
D. **Values identification.** The perspectives, procedures, and rationale used to interpret the findings should be carefully described so that the bases for value judgments are clear.
E. **Report clarity.** Evaluation reports should clearly describe the program being evaluated, including its context and the purposes, procedures, and findings of the evaluation so that essential information is provided and easily understood.
F. **Report timeliness and dissemination.** Substantial interim findings and evaluation reports should be disseminated to intended users so that they can be used in a timely fashion.
G. **Evaluation impact.** Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders to increase the likelihood of the evaluation being used.

Source: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

BOX 14. Feasibility standards

The following feasibility standards ensure that an evaluation will be realistic, prudent, diplomatic, and frugal:

A. **Practical procedures.** Evaluation procedures should be practical while needed information is being obtained to keep disruption to a minimum.
B. **Political viability.** During planning and conduct of the evaluation, consideration should be given to the varied positions of interest groups so that their cooperation can be obtained and possible attempts by any group to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.
C. **Cost-effectiveness.** The evaluation should be efficient and produce valuable information to justify expended resources.

Source: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

Standard 4: Accuracy

Accuracy standards ensure that the evaluation produces findings that are considered correct. Twelve accuracy standards (Box 16) include such items as describing the program and its context; articulating in detail the purpose and methods of the evaluation; employing systematic procedures to gather valid and reliable information; applying appropriate qualitative or quantitative methods during analysis and synthesis; and producing impartial reports containing conclusions that are justified.

The steps and standards are used together throughout the evaluation process. For each step, a subset of relevant standards should be considered (Box 17).

BOX 15. Propriety standards

The following propriety standards ensure that an evaluation will be conducted legally, ethically, and with regard for the welfare of those involved in the evaluation as well as those affected by its results:

A. **Service orientation.** The evaluation should be designed to assist organizations in addressing and serving effectively the needs of the targeted participants.
B. **Formal agreements.** All principal parties involved in an evaluation should agree in writing to their obligations (i.e., what is to be done, how, by whom, and when) so that each must adhere to the conditions of the agreement or renegotiate it.
C. **Rights of human subjects.** The evaluation should be designed and conducted in a manner that respects and protects the rights and welfare of human subjects.
D. **Human interactions.** Evaluators should interact respectfully with other persons associated with an evaluation, so that participants are not threatened or harmed.
E. **Complete and fair assessment.** The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program so that strengths can be enhanced and problem areas addressed.
F. **Disclosure of findings.** The principal parties to an evaluation should ensure that the full evaluation findings with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results.
G. **Conflict of interest.** Conflict of interest should be handled openly and honestly so that the evaluation processes and results are not compromised.
H. **Fiscal responsibility.** The evaluator's allocation and expenditure of resources should reflect sound accountability procedures by being prudent and ethically responsible, so that expenditures are accountable and appropriate.

Source: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

BOX 16. Accuracy standards

The following accuracy standards ensure that an evaluation will convey technically adequate information regarding the determining features of merit of the program:

A. **Program documentation.** The program being evaluated should be documented clearly and accurately.
B. **Context analysis.** The context in which the program exists should be examined in enough detail to identify probable influences on the program.
C. **Described purposes and procedures.** The purposes and procedures of the evaluation should be monitored and described in enough detail to identify and assess them.
D. **Defensible information sources.** Sources of information used in a program evaluation should be described in enough detail to assess the adequacy of the information.
E. **Valid information.** Information-gathering procedures should be developed and implemented to ensure a valid interpretation for the intended use.
F. **Reliable information.** Information-gathering procedures should be developed and implemented to ensure sufficiently reliable information for the intended use.
G. **Systematic information.** Information collected, processed, and reported in an evaluation should be systematically reviewed and any errors corrected.
H. **Analysis of quantitative information.** Quantitative information should be analyzed appropriately and systematically so that evaluation questions are answered effectively.
I. **Analysis of qualitative information.** Qualitative information should be analyzed appropriately and systematically to answer evaluation questions effectively.
J. **Justified conclusions.** Conclusions reached should be explicitly justified for stakeholders' assessment.
K. **Impartial reporting.** Reporting procedures should guard against the distortion caused by personal feelings and biases of any party involved in the evaluation to reflect the findings fairly.
L. **Metaevaluation.** The evaluation should be formatively and summatively evaluated against these and other pertinent standards to guide its conduct appropriately and, on completion, to enable close examination of its strengths and weaknesses by stakeholders.

Source: Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

BOX 17. Cross-reference of steps and relevant standards

Steps in Evaluation Practice	Relevant Standards	Group/ Box No.-Item
Engaging stakeholders	Stakeholder identification	Utility/13-A
	Evaluator credibility	Utility/13-B
	Formal agreements	Propriety/15-B
	Rights of human subjects	Propriety/15-C
	Human interactions	Propriety/15-D
	Conflict of interest	Propriety/15-G
	Metaevaluation	Accuracy/16-L
Describing the program	Complete and fair assessment	Propriety/15-C
	Program documentation	Accuracy/16-A
	Context analysis	Accuracy/16-B
	Metaevaluation	Accuracy/16-L
Focusing the evaluation design	Evaluation impact	Utility/13-G
	Practical procedures	Feasibility/14-A
	Political viability	Feasibility/14-B
	Cost effectiveness	Feasibility/14-C
	Service orientation	Propriety/15-A
	Complete and fair assessment	Propriety/15-E
	Fiscal responsibility	Propriety/15-H
	Described purposes and procedures	Accuracy/16-C
	Metaevaluation	Accuracy/16-C
Gathering credible evidence	Information scope and selection	Utility/13-C
	Defensible information sources	Accuracy/16-D
	Valid information	Accuracy/16-E
	Reliable information	Accuracy/16-F
	Systematic information	Accuracy/16-G
	Metaevaluation	Accuracy/16-L
Justifying conclusions	Values identification	Utility/13-D
	Analysis of quantitative information	Accuracy/16-H
	Analysis of qualitative information	Accuracy/16-I
	Justified conclusions	Accuracy/16-J
	Metaevaluation	Accuracy/16-L
Ensuring use and sharing lessons learned	Evaluator credibility	Utility/13-B
	Report clarity	Utility/13-E
	Report timeliness and dissemination	Utility/13-F
	Evaluation impact	Utility/13-G
	Disclosure of findings	Propriety/15-F
	Impartial reporting	Accuracy/16-K
	Metaevaluation	Accuracy/16-L

APPLYING THE FRAMEWORK

Conducting Optimal Evaluations

Public health professionals can no longer question whether to evaluate their programs; instead, the appropriate questions are

- What is the best way to evaluate?
- What is being learned from the evaluation? And,

- How will lessons learned from evaluations be used to make public health efforts more effective and accountable?

The framework for program evaluation helps answer these questions by guiding its users in selecting evaluation strategies that are useful, feasible, ethical, and accurate. To use the recommended framework in a specific program context requires practice, which builds skill in both the science and art of program evaluation. When applying the framework, the challenge is to devise an optimal — as opposed to an ideal — strategy. An optimal strategy is one that accomplishes each step in the framework in a way that accommodates the program context and meets or exceeds all relevant standards. CDC's evaluations of human immunodeficiency virus prevention efforts, including school-based programs, provide examples of optimal strategies for national-, state-, and local-level evaluation (*97,98*).

Assembling an Evaluation Team

Harnessing and focusing the efforts of a collaborative group is one approach to conducting an optimal evaluation (*24,25*). A team approach can succeed when a small group of carefully selected persons decides what the evaluation must accomplish and pools resources to implement the plan. Stakeholders might have varying levels of involvement on the team that correspond to their own perspectives, skills, and concerns. A leader must be designated to coordinate the team and maintain continuity throughout the process; thereafter, the steps in evaluation practice guide the selection of team members. For example,

- Those who are diplomatic and have diverse networks can engage other stakeholders and maintain involvement.
- When describing the program, persons are needed who understand the program's history, purpose, and practical operation in the field. In addition, those with group facilitation skills might be asked to help elicit unspoken expectations regarding the program and to expose hidden values that partners bring to the effort. Such facilitators can also help the stakeholders create logic models that describe the program and clarify its pattern of relationships between means and ends.
- Decision makers and others who guide program direction can help focus the evaluation design on questions that address specific users and uses. They can also set logistic parameters for the evaluation's scope, time line, and deliverables.
- Scientists, particularly social and behavioral scientists, can bring expertise to the development of evaluation questions, methods, and evidence gathering strategies. They can also help analyze how a program operates in its organizational or community context.
- Trusted persons who have no particular stake in the evaluation can ensure that participants' values are treated fairly when applying standards, interpreting facts, and reaching justified conclusions.

- Advocates, clear communicators, creative thinkers, and members of the power structure can help ensure that lessons learned from the evaluation influence future decision-making regarding program strategy.

All organizations, even those that are able to find evaluation team members within their own agency, should collaborate with partners and take advantage of community resources when assembling an evaluation team. This strategy increases the available resources and enhances the evaluation's credibility. Furthermore, a diverse team of engaged stakeholders has a greater probability of conducting a culturally competent evaluation (i.e., one that understands and is sensitive to the persons, conditions, and contexts associated with the program) (*99,100*). Although challenging for the coordinator and the participants, the collaborative approach is practical because of the benefits it brings (e.g., reduces suspicion and fear, increases awareness and commitment, increases the possibility of achieving objectives, broadens knowledge base, teaches evaluation skills, strengthens partnerships, increases the possibility that findings will be used, and allows for differing perspectives) (*8,24*).

Addressing Common Concerns

Common concerns regarding program evaluation are clarified by using this framework. Evaluations might not be undertaken because they are misperceived as having to be costly. However, the expense of an evaluation is relative; the cost depends on the questions being asked and the level of precision desired for the answers. A simple, low-cost evaluation can deliver valuable results.

Rather than discounting evaluations as time-consuming and tangential to program operations (e.g., left to the end of a program's project period), the framework encourages conducting evaluations from the beginning that are timed strategically to provide necessary feedback to guide action. This makes integrating evaluation with program practice possible.

Another concern centers on the perceived technical demands of designing and conducting an evaluation. Although circumstances exist where controlled environments and elaborate analytic techniques are needed, most public health program evaluations do not require such methods. Instead, the practical approach endorsed by this framework focuses on questions that will improve the program by using context-sensitive methods and analytic techniques that summarize accurately the meaning of qualitative and quantitative information.

Finally, the prospect of evaluation troubles some program staff because they perceive evaluation methods as punitive, exclusionary, or adversarial. The framework encourages an evaluation approach that is designed to be helpful and engages all interested stakeholders in a process that welcomes their participation. Sanctions to be applied, if any, should not result from discovering negative findings, but from failing to use the learning to change for greater effectiveness (*10*).

EVALUATION TRENDS

Interest in program improvement and accountability continues to grow in government, private, and nonprofit sectors. The Government Performance and Results Act

requires federal agencies to set performance goals and to measure annual results. Nonprofit donor organizations (e.g., United Way) have integrated evaluation into their program activities and now require that grant recipients measure program outcomes (*30*). Public-health-oriented foundations (e.g., W.K. Kellogg Foundation) have also begun to emphasize the role of evaluation in their programming (*24*). Innovative approaches to staffing program evaluations have also emerged. For example, the American Cancer Society (ACS) Collaborative Evaluation Fellows Project links students and faculty in 17 schools of public health with the ACS national and regional offices to evaluate local cancer control programs (*101*). These activities across public and private sectors reflect a collective investment in building evaluation capacity for improving performance and being accountable for achieving public health results.

Investments in evaluation capacity are made to improve program quality and effectiveness. One of the best examples of the beneficial effects of conducting evaluations is the Malcolm Baldridge National Quality Award Program (*102*).* Evidence demonstrates that the evaluative processes required to win the Baldridge Award have helped American businesses outperform their competitors (*103*). Now these same effects on quality and performance are being translated to the health and human service sector. Recently, Baldridge Award criteria were developed for judging the excellence of health care organizations (*104*). This extension to the health-care industry illustrates the critical role for evaluation in achieving health and human service objectives. Likewise, the framework for program evaluation was developed to help integrate evaluation into the corporate culture of public health and fulfill CDC's operating principles for public health practice (*1,2*).

Building evaluation capacity throughout the public health workforce is a goal also shared by the Public Health Functions Steering Committee. Chaired by the U.S. Surgeon General, this committee identified core competencies for evaluation as essential for the public health workforce of the twenty first century (*105*). With its focus on making evaluation accessible to all program staff and stakeholders, the framework helps to promote evaluation literacy and competency among all public health professionals.

SUMMARY

Evaluation is the only way to separate programs that promote health and prevent injury, disease, or disability from those that do not; it is a driving force for planning effective public health strategies, improving existing programs, and demonstrating the results of resource investments. Evaluation also focuses attention on the common purpose of public health programs and asks whether the magnitude of investment matches the tasks to be accomplished (*95*).

The recommended framework is both a synthesis of existing evaluation practices and a standard for further improvement. It supports a practical approach to evaluation that is based on steps and standards applicable in public health settings. Because the framework is purposefully general, it provides a guide for designing and conducting

*The Malcolm Baldridge National Quality Improvement Act of 1987 (Public Law 100-107) established a public-private partnership focused on encouraging American business and other organizations to practice effective quality management. The annual award process, which involves external review as well as self-assessment against Criteria for Performance Excellence, provides a proven course for organizations to improve significantly the quality of their goods and services.

specific evaluation projects across many different program areas. In addition, the framework can be used as a template to create or enhance program-specific evaluation guidelines that further operationalize the steps and standards in ways that are appropriate for each program (*20,96,106–112*). Thus, the recommended framework is one of several tools that CDC can use with its partners to improve and account for their health promotion and disease or injury prevention work.

ADDITIONAL INFORMATION

Sources of additional information are available for those who wish to begin applying the framework presented in this report or who wish to enhance their understanding of program evaluation. In particular, the following resources are recommended:

- "Practical Evaluation of Public Health Programs" (course no. VC0017) is a 5-hr distance-learning course that also uses the framework presented in this report. Developed through CDC's Public Health Training Network (PHTN) (*8*), the course consists of two videotapes and a workbook, which can be used by individuals for self-study or by small groups with optional activities. Continuing education credit is available for this course. Additional information is available at the PHTN website at <http://www.cdc.gov/phtn> or by calling, toll-free, 800-41-TRAIN (800-418-7246). Also, course materials can be purchased from the Public Health Foundation by calling, toll-free, 877-252-1200, or using the on-line order form at <http://bookstore.phf.org/prod41.htm>. For informational purposes, the workbook can be viewed over the Internet at <http://www.cdc.gov/eval/workbook.pdf>.
- The Community Toolbox (CTB) is an Internet resource for health promotion and community development that contains information regarding how to conduct public health work and social change on a community level. Because they consider program evaluation to be a critical part of successful community-based health promotion, the CTB team used the framework for program evaluation to create a unique gateway to evaluation-related ideas and tools. This gateway can be accessed at <http://ctb.lsi.ukans.edu/ctb/c30/ProgEval.html>.
- The CDC Evaluation Working Group has compiled a list of additional resources for program evaluation. These resources address such topics as a) ethics, principles, and standards for program evaluation; b) evaluation-related organizations, societies, foundations, and associations; c) journals and on-line publications; d) step-by-step evaluation manuals; e) resources for developing logic models; f) planning- and performance-improvement tools; and g) evaluation-related publications. This list of resources can be obtained through the Working Group's website at <http://www.cdc.gov/eval/index.htm> or by sending an electronic message to <eval@cdc.gov>.

References

1. Dyal WW. Ten organizational practices of public health: a historical perspective. Am J Prev Med 1995;11(6)Suppl 2:6–8.
2. Koplan JP. CDC sets millennium priorities. US Medicine 1999;4–7.

3. Scriven M. Minimalist theory of evaluation: the least theory that practice requires. American Journal of Evaluation 1998;19:57–70.
4. Shadish WR, Cook TD, Leviton LC. Foundations of program evaluation: theories of practice. Newbury Park, CA: Sage Publications, 1991.
5. Weiss CH. Evaluation: methods for studying programs and policies. 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1998.
6. Worthen BR, Sanders JR, Fitzpatrick, JL. Program evaluation: alternative approaches and practical guidelines. 2nd ed. New York, NY: Longman, 1996.
7. Patton MQ. Utilization-focused evaluation: the new century text. 3rd ed. Thousand Oaks, CA: Sage Publications, 1997.
8. CDC. Practical evaluation of public health programs. Atlanta, GA: US Department of Health and Human Services, CDC, Public Health Training Network, 1998; PHTN course no. VC-0017.
9. Love A. Internal evaluation: building organizations from within. Applied social research, vol 24. Newbury Park, CA: Sage Publications, 1991.
10. Sanders JR. Uses of evaluation as a means toward organizational effectiveness. In: Gray ST, ed. Leadership IS: a vision of evaluation; a report of learnings from Independent Sector's work on evaluation. Washington, DC: Independent Sector, 1993.
11. Shadish WR. Evaluation theory is who we are. American Journal of Evaluation 1998;19(1):1–19.
12. Joint Committee on Standards for Educational Evaluation. Program evaluation standards: how to assess evaluations of educational programs. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
13. Joint Committee on Standards for Educational Evaluation. Personnel evaluation standards: how to assess systems for evaluating educators. Newbury Park, CA: Sage Publications, 1988.
14. CDC. Principles of community engagement. Atlanta, GA: CDC, Public Health Practice Program Office, 1997.
15. Mertens DM. Inclusive evaluation: implications of transformative theory for evaluation. American Journal of Evaluation 1999;20(1):1–14.
16. Connell JP, Kubisch AC. Applying a theory of change approach to the evaluation of comprehensive community initiatives: progress, prospects, and problems. In: Fulbright-Anderson K, Kubisch AC, Connell JP, eds. New approaches to evaluating community initiatives: theory, measurement, and analysis. Washington, DC: Aspen Institute, 1998.
17. Chen HT. Theory driven evaluations. Newbury Park, CA: Sage Publications, 1990.
18. U.S. General Accounting Office. Managing for results: measuring program results that are under limited federal control. Washington, DC: U.S. General Accounting Office 1998; publication no. GAO/GGD-99-16.
19. Haddix AC, Teutsch SM, Shaffer PA, Duñet DO, eds. Prevention effectiveness: a guide to decision analysis and economic evaluation. New York, NY: Oxford University Press, 1996.
20. CDC. Framework for assessing the effectiveness of disease and injury prevention. MMWR 1992;41(No. RR-3):1–13.
21. Gold ME, Russell LB, Siegel JE, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press, 1996.
22. Eoyang GH, Berkas T. Evaluation in a complex adaptive system. In: Lissack M, Gunz H, eds. Managing complexity in organizations. Westport, CT: Quorum Books, 1999.
23. Rush B, Ogbourne A. Program logic models: expanding their role and structure for program planning and evaluation. Canadian Journal of Program Evaluation 1991;6(2):95–106.
24. W.K. Kellogg Foundation. W.K. Foundation Evaluation Handbook. Battle Creek, MI: W.K. Kellogg Foundation, 1998.
25. Taylor-Powell E, Rossing B, Geran J. Evaluating collaboratives: reaching the potential. Madison, WI: University of Wisconsin—Cooperative Extension, 1998.
26. Poreteous NL, Sheldrick BJ, Stewart PJ. Program evaluation tool kit: a blueprint for public health management. Ottawa, Ontario, Canada: Ottawa-Carleton Health Department, Public Health Research, Education and Development Program, 1997.
27. Weiss CH. Nothing as practical as a good theory: exploring theory-based evaluation for comprehensive community initiatives for families and children. In: Connell JP, Kubisch AC, Schorr LB, Weiss CH, eds. New approaches to evaluating community initiatives: concepts, methods, and contexts. Washington, DC: Aspen Institute, 1995.

28. Weiss CH. How can theory-based evaluation make greater headway? Evaluation Review 1997;21:501–24.
29. Wong-Reiger D, David L. Using program logic models to plan and evaluate education and prevention programs. In: Love AJ, ed. Evaluation methods sourcebook II. Ottawa, Ontario, Canada: Canadian Evaluation Society, 1995.
30. United Way of America. Measuring program outcomes: a practical approach. Alexandria, VA: United Way of America, 1996.
31. Moyer A, Verhovsek H, Wilson VL. Facilitating the shift to population-based public health programs: innovation through the use of framework and logic model tools. Can J Public Health 1997;88(2):95–8.
32. McLaughlin JA, Jordan GB. Logic models: a tool for telling your program's performance story. Evaluation and Program Planning 1999;22(1):65–72.
33. McEwan KL, Bigelow DA. Using a logic model to focus health services on population health goals. Canadian Journal of Program Evaluation 1997;12(1):167–74.
34. Julian D. Utilization of the logic model as a system level planning and evaluation device. Evaluation and Program Planning 1997;20(3):251–7.
35. Lipsey MW. Theory as method: small theories of treatments. New Directions for Program Evaluation 1993;57:5–38.
36. U.S. General Accounting Office. Designing evaluations. Washington, DC: U.S. General Accounting Office, 1991; publication no. GAO/PEMD-10.1.4.
37. Taylor-Powell E, Steele S, Douglah M. Planning a program evaluation. Madison, WI: University of Wisconsin Cooperative Extension, 1996.
38. Fetterman DM, Kaftarian SJ, Wandersman A, eds. Empowerment evaluation: knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage Publications, 1996.
39. Preskill HS, Torres RT. Evaluative inquiry for learning in organizations. Thousand Oaks, CA: Sage Publications, 1999.
40. Patton MQ. Toward distinguishing empowerment evaluation and placing it in a larger context. Evaluation Practice 1997;18(2):147–63.
41. Wandersman A, Morrissey E, Davino K, et al. Comprehensive quality programming and accountability: eight essential strategies for implementing successful prevention programs. Journal of Primary Prevention 1998;19(1):3–30.
42. Cousins JB, Whitmore E. Framing participatory evaluation. In: Whitmore E, ed. Understanding and practicing participatory evaluation, vol 80. San Francisco, CA: Jossey-Bass, 1998:5–24.
43. Bickman L, Rog DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998.
44. Cook TD, Reichardt CS, eds. Qualitative and quantitative methods in evaluation research. Sage research progress series in evaluation, vol 1. Beverly Hills, CA: Sage Publications, 1979.
45. Patton MQ. Qualitative evaluation and research methods. 2nd ed. Newbury Park, CA: Sage Publications, 1990.
46. Posavac EJ, Carey RG. Program evaluation: methods and case studies. Englewood Cliffs, NJ: Prentice-Hall, 1980.
47. Rossi PH, Freeman HE, Lipsey MW. Evaluation: a systematic approach. 6th ed. Thousand Oaks, CA: Sage Publications, 1999.
48. Trochim WMK. Research methods knowledge base [on-line textbook]. 2nd ed. 1999. Available at <http://trochim.human.cornell.edu >. Accessed June 1999.
49. Boruch RF. Randomized controlled experiments for evaluation and planning. In: Bickman L, Rog DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:161–92.
50. Cook TD, Campbell DT. Quasi-experimentation: design and analysis issues for field settings. Boston, MA: Houghton Mifflin, 1979.
51. Reichardt CS, Mark MM. Quasi-experimentation. In: Bickman L, Rob DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:193–228.
52. Yin RK. Case study research: design and methods. 2nd ed. Applied Social Research Methods Series, vol 5. Thousand Oaks, CA: Sage Publications, 1994.
53. Yin RK. Abridged version of case study research: design and method. In: Bickman L, Rob DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:229–60.

54. U.S. General Accounting Office. Case study evaluations. Washington, DC: U.S. General Accounting Office, 1990; publication no. GAO/PEMD-91-10.1.9.
55. McQueen DV, Anderson LM. What counts as evidence? Issues and debates on evidence relevant to the evaluation of community health promotion programs. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. Evaluation in health promotion: principles and perspectives. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
56. Green JC, Caracelli V, eds. Advances in mixed-method evaluation: the challenges and benefits of integrating diverse paradigms. New Directions for Program Evaluation, vol 74; San Francisco, CA: Jossey-Bass 1997.
57. Frechtling J, Sharp L. User-friendly handbook for mixed method evaluations. Arlington, VA: National Science Foundation, 1997.
58. U.S. General Accounting Office. Evaluation synthesis. Washington, DC: U.S. General Accounting Office, 1992. publication no. GAO/PEMD-10.1.2.
59. Newman DL, Brown RD. Applied ethics for program evaluation. Thousand Oaks, CA: Sage Publications, 1996.
60. Fitzpatrick JL, Morris M, eds. Current and emerging ethical challenges in evaluation. New Directions for Program Evaluation, vol 82; San Francisco, CA: Jossey-Bass 1999.
61. Newcomer K. Using statistics appropriately. In: Wholey J, Hatry H, Newcomer K, eds. Handbook of practical program evaluation. San Francisco, CA: Jossey-Bass, 1994.
62. Basch CE, Silepcevich EM, Gold RS, Duncan DF, Kolbe LJ. Avoiding type III errors in health education program evaluation: a case study. Health Education Quarterly 1985;12(4):315–31.
63. Perrin EB, Koshel JJ, eds. Assessment of performance measures for public health, substance abuse, and mental health. Washington, DC: National Academy Press, 1997.
64. Innes JE. Knowledge and public policy: the search for meaningful indicators. 2nd expanded ed. New Brunswick, NJ: Transaction Publishers, 1990.
65. McRae D Jr. Policy indicators: links between social science and public debate. Chapel Hill, NC: University of North Carolina Press, 1985.
66. Institute of Medicine. Improving health in the community: a role for performance monitoring. Durch, JS, Bailey LA, Stoto MA, eds. Washington, DC: National Academy Press, 1997.
67. Eddy DM. Performance measurement: problems and solutions. Health Aff 1998;17(4):7–25.
68. Harvard Family Research Project. Performance measurement. Evaluation Exchange 1998;4(1):1–15.
69. Rugg D. New activities and use of indicators in the evaluation of HIV prevention efforts at CDC. National Alliance of State and Territorial AIDS Directors (NASTAD) HIV Prevention Community Planning Bulletin;1997:2–3.
70. Center for Substance Abuse Prevention. Measurements in prevention: a manual on selecting and using instruments to evaluate prevention programs. CSAP technical report no. 8 (SMA)93-2041. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, 1993.
71. Goodman RM, Speers MA, McLeroy K, et al. Identifying and defining the dimensions of community capacity to provide a basis for measurement. Health Educ Behav 1998;25(3):258–78.
72. Phillips JJ. Handbook of training evaluation and measurement methods. 3rd ed. Houston, TX: Gulf Publishing Company, 1997.
73. Knauft EB. What Independent Sector learned from an evaluation of its own hard-to-measure programs. In: Gray ST, ed. Leadership IS: a vision of evaluation; a report of learnings from Independent Sector's work on evaluation. Washington, DC: Independent Sector, 1993.
74. Perrin B. Effective use and misuse of performance measurement. American Journal of Evaluation 1998;19(3):367–79.
75. U.S. General Accounting Office. Using statistical sampling. Washington, DC: U.S. General Accounting Office, 1992; publication no. GAO/PEMD-10.1.6.
76. Henry GT. Practical sampling. In: Bickman L, Rog DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:101–26.
77. Johnson JC. Selecting ethnographic informants. Qualitative Research Methods Series, vol 22; Thousand Oaks, CA: Sage Publications, 1991,
78. Steckler A, McLeroy KR, Goodman RM, Bird ST, McCormick L. Toward integrating qualitative and quantitative methods: an introduction. Health Education Quarterly 1992;19(1):191–8.

79. Greene JC. Qualitative program evaluation: practice and promise. In: Denzin NK, Lincoln YS, eds. Handbook of qualitative research. Thousand Oaks, CA: Sage Publications, 1994.
80. de Vries H, Weijts W, Dijkstra M, Kok G. Utilization of qualitative and quantitative data for health education program planning, implementation, and evaluation: a spiral approach. Health Education Quarterly 1992;19:101–15.
81. U.S. General Accounting Office. Prospective evaluation methods: the prospective evaluation synthesis. Washington, DC: U.S. General Accounting Office, 1990; publication no. GAO/PEMD-10.1.10.
82. Lipsey MW. What can you build with thousands of bricks? Musings on the cumulation of knowledge in program evaluation. New Directions for Evaluation 1997;76:7–23.
83. Lipsy MW. Design sensitivity: statistical power for applied experimental research. In: Bickman L, Rob DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:39–68.
84. Sieber JE. Planning ethically responsible research. In: Bickman L, Rob DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:127–56.
85. Miles MB, Huberman AM. Qualitative data analysis: an expanded sourcebook. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
86. Hennessy M. Evaluation. In: Stroup D, Teutsch S., eds. Statistics in Public Health. New York, NY: Oxford University Press, 1998:193–219.
87. Henry GT. Graphing data. In: Bickman L, Rog DJ, eds. Handbook of applied social research methods. Thousand Oaks, CA: Sage Publications, 1998:527–56.
88. Weick KE. Sensemaking in organizations. Thousand Oaks, CA: Sage Publications, 1995.
89. Rogers PJ, Hough G. Improving the effectiveness of evaluations: making the link to organizational theory. Evaluation and Program Planning 1995;18(4):321–32.
90. Adler M, Ziglio E. Gazing into the oracle: the Delphi method and its application to social policy and public health. Bristol, PA: Jessica Kingsley Publishers, 1996.
91. Torres RT, Preskill HS, Piontek ME. Evaluation strategies for communicating and reporting: enhancing learning in organizations. Thousand Oaks, CA: Sage Publications, 1996.
92. Zaltman G, Barabba VP. Hearing the voice of the market: competitive advantage through creative use of market information. Boston, MA: Harvard Business School Press, 1990.
93. Shulha LM, Cousins JB. Evaluation use: theory, research, and practice since 1986. Evaluation Practice 1997;18(3):195–208.
94. Weiss CH. Have we learned anything new about the use of evaluation? American Journal of Evaluation 1998;19(1):21–33.
95. Schorr LB. Common purpose: strengthening families and neighborhoods to rebuild America. New York, NY: Doubleday, Anchor Books, 1997.
96. Fawcett SB, Paine-Andrews A, Francisco VT, et al. Evaluating community initiatives for health and development. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. Evaluation in health promotion: principles and perspectives. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
97. Collins J, Rugg D, Kann L, Banspach S, and Kolbe L. Evaluating a national program of school-based HIV prevention. Evaluation and Program Planning 1996;19(3):209–18.
98. Rugg D, Buehler J, Renaud M, et al. Evaluating HIV prevention: a framework for national, state, and local levels. American Journal of Evaluation 1999:20(1):35–56.
99. Greene J, Lincoln Y, Mathison S, Mertens DM, Ryan K. Advantages and challenges of using inclusive evaluation approaches in evaluation practice. American Journal of Evaluation 1998;19(1):101–22.
100. Smith NL. The context of investigations in cross-cultural evaluations. Studies in Educational Evaluation 1991;17:3–21.
101. Bonnet, DG. An evaluation of year 1 of the American Cancer Society's collaborative evaluation fellows project. Indianapolis, IN: D. Bonnet Associates, 1999.
102. National Institute of Standards and Technology. National Quality Program. National Institute of Standards and Technology, 1999. Available at <http://www.quality.nist.gov>. Accessed June 1999.
103. National Institute of Standards and Technology. Baldridge index outperforms S&P 500 for fifth year. National Institute of Standards and Technology, National Quality Program, 1999. Available at <http://www.nist.gov/public_affairs/releases/n99-02.htm>. Accessed June 1999.

104. National Institute of Standards and Technology. Health care criteria for performance excellence. Gaithersburg, MD: National Institute of Standards and Technology, Baldrige National Quality Program, 1999. Also available at <http://www.quality.nist.gov>. Accessed June 1999.
105. Public Health Service. Public health workforce: an agenda for the 21st century: a report of the public health functions project. Washington, DC: US Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 1997.
106. CDC. Handbook for evaluating HIV education. Atlanta, GA: US Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, Division of Adolescent and School Health, 1993.
107. Thompson, NJ, McClintock HO. Demonstrating your program's worth: a primer on evaluation for programs to prevent unintentional injury. Atlanta: CDC, National Center for Injury Prevention and Control, 1998.
108. CDC. Guidelines for evaluating surveillance systems. MMWR 1988;37(No. SS-5):1–18.
109. Fawcett SB, Sterling TD, Paine-Andrews A, et al. Evaluating community efforts to prevent cardiovascular diseases. Atlanta, GA: U.S. Department of Health and Human Services, CDC, National Center for Chronic Disease Prevention and Health Promotion, 1995.
110. Rootman I, Goodstadt M, Hyndman B, et al., eds. Evaluation in health promotion: principles and perspectives. Copenhagen, Denmark. World Health Organization (Euro), 1999 (In press).
111. CDCynergy (Version 1.0) [CD ROM]. Atlanta, GA: U.S. Department of Health and Human Services, CDC, Office of Communication, 1998.
112. Linney JA, Wandersman A. Prevention plus III: assessing alcohol and other drug prevention programs at the school and community level: a four-step guide to useful program assessment. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Alcohol, Drug Abuse, and Mental Health Administration, Office for Substance Abuse Prevention, 1991. publication no. (ADM)91-1817.

The *Morbidity and Mortality Weekly Report (MMWR)* Series is prepared by the Centers for Disease Control and Prevention (CDC) and is available free of charge in electronic format and on a paid subscription basis for paper copy. To receive an electronic copy on Friday of each week, send an e-mail message to *listserv@listserv.cdc.gov*. The body content should read *SUBscribe mmwr-toc*. Electronic copy also is available from CDC's World-Wide Web server at *http://www.cdc.gov/* or from CDC's file transfer protocol server at *ftp.cdc.gov*. To subscribe for paper copy, contact Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402; telephone (202) 512-1800.

Data in the weekly *MMWR* are provisional, based on weekly reports to CDC by state health departments. The reporting week concludes at close of business on Friday; compiled data on a national basis are officially released to the public on the following Friday. Address inquiries about the *MMWR* Series, including material to be considered for publication, to: Editor, *MMWR* Series, Mailstop C-08, CDC, 1600 Clifton Rd., N.E., Atlanta, GA 30333; telephone (888) 232-3228.

☆**U.S. Government Printing Office: 1999-733-228/08020 Region IV**