3 Focus the Evaluation Design

Now that you and your stakeholders have a clear understanding of your program, your evaluation team will need to focus the evaluation. The evaluation team must decide the purpose of the evaluation and the questions it wants answered. A typical approach to evaluation in public health is to design data-collection systems that monitor progress toward meeting a program's process and outcome objectives. Initially, you may not be able to collect baseline data and track progress toward all of your objectives. However, it is important to remember that baseline data are valuable for planning and evaluation and should be collected if possible. Rather than trying to answer every question that various stakeholders may pose, the evaluation team should focus on those it determines to be the most important questions about your program. A focused evaluation requires "advance planning about where the evaluation is headed and what steps will be taken to get there."[3]

Having a focused evaluation makes it easier to conduct a quality evaluation. The design should outline which questions you are investigating, the process you will follow, what will be measured, what methods will be used, who will perform each activity (including analysis and interpretation), what you will do with the information once it is collected, and how the results will be disseminated.

Process evaluation

Process evaluations are used to document how well a program has been implemented; they are conducted periodically throughout the duration of a program. This type of evaluation is used to examine the operations of a program, including which activities are taking place, who is conducting the activities, and who is reached through the activities. Process evaluations assess whether inputs or resources have been allocated or mobilized and whether activities are being implemented as planned. They identify program strengths, weaknesses, and areas that need

improvement. Following are examples of the type of tangible program indicators measured by process evaluation:

- The locale where services or programs are provided (e.g., rural, urban).
- The number of people receiving services.
- The economic status and racial/ethnic background of people receiving services.
- The quality of services.
- The actual events that occur while the services are delivered.
- The amount of money the project is using.
- The direct and in-kind funding for services.
- The staffing for services or programs.
- The number of activities and meetings.
- The number of training sessions conducted.

A process evaluation of a counter-marketing campaign to reduce the number of young people who start smoking might answer questions such as these:

- Has a workgroup been formed and is it meeting regularly?
- Are any key individuals or organizations missing from the workgroup?
- Was the counter-marketing campaign designed on schedule?
- Have the campaign products (posters, billboard, radio and television spots) been pretested?
- Are project activities being implemented on schedule?
- What barriers have been encountered?
- Who is the campaign's target audience and how well are they being reached?
- How many advertisements are actually running? When and where?
- Where are the posters/billboards located?
- What is the estimated number of people who see or hear the advertisements?
- How might the action plan be improved on the basis of evaluation findings?

Process evaluations can also assess issues related to program services. For example, they can determine the—

- Availability and use of tobacco-use treatment services.
- Implementation of smoking prevention programs in schools and the community.
- Accessibility of resource centers and materials.
- Amount of technical support and training provided to grantees or staff.
- Amount of technical support and training needed by grantees or staff.
- Number of calls to a quitline.
- Use of the quitline by various racial/ethnic groups.
- Extent of insurance coverage for tobacco-use treatment.
- Percentage of primary care physicians who give advice and assistance on quitting.
- Number of health care systems that have implemented tobacco-use reminder systems.
- Use of Food and Drug Administration (FDA)-approved medications by Medicaid recipients.

These are straightforward questions; monitoring them throughout the duration of your program ensures that the project is implemented as planned and is reaching the intended audience.

Outcome evaluation

Outcome evaluations are used to assess the impact of a program on the stated short-term, intermediate, and long-term objectives. This type of evaluation assesses what has occurred because of the program and whether the program has achieved its outcome objectives. Outcome evaluations should be conducted only when the program is mature enough to produce the intended outcome.

Outcome evaluations can measure the following:

- Changes in people's attitude toward, and beliefs about, tobacco, their awareness of and support for your program, and their perception of how well tobacco-related policy is being enforced.

- Changes in intended and actual tobacco-related behaviors.
- Changes in the environment, such as changes in public and private policies, in formal and informal enforcement of minors' access and nonsmoking regulations, and in the influence of pro-tobacco forces.
- Changes in populations, such as in the average age at which people begin smoking, per capita consumption of cigarettes, and smoking prevalence.
- Changes in trends in morbidity and mortality.

In this manual, program outcomes are divided into three levels: short-term, intermediate, and long-term. Decisions as to whether a particular outcome is short-term, intermediate, or long-term depend on the purpose of the program and the time needed for the change to occur. For example, there are no strict guidelines for whether a policy change is a short-term or an intermediate outcome; it could also be thought of as a process measure.

Similarly, changes in per capita consumption could be considered an intermediate or a long-term outcome. Whether outcomes are considered short- or long-term is less important than whether sound logic underlies the program. Do the short-term outcomes lead logically to the intermediate outcomes? Do the intermediate outcomes lead logically to the long-term outcomes? Is adequate time allowed to reasonably expect to see an effect?

Short-term outcomes are the immediate or early results of the program. Short-term outcomes may be changes in knowledge, attitudes, and skills. For example, in a program with the goal of reducing children's exposure to ETS, a short-term outcome might be having parents who smoke show increased knowledge about the danger of smoking around children.

Intermediate outcomes reflect further progress in reaching a program goal. Intermediate outcomes link short-term outcomes with long-term outcomes. Intermediate outcomes may be changes in individual behaviors, social norms, or the environment. An intermediate outcome in the program described in the previous paragraph might be that the parents no longer smoke around their children.

Long-term outcomes reflect the ultimate goal of the program. The long-term outcome in the previously described program would be decreased morbidity from children's exposure to ETS.

For a tobacco control program with the goal of reducing the number of young people who start smoking through a counter-marketing campaign, an outcome evaluation might examine whether the targeted young people exhibit—

- Increased knowledge and awareness of the dangers of smoking (short-term outcome).
- Changes in tobacco-related attitudes and beliefs (intermediate outcome).
- Changes in tobacco-related behavior (long-term outcome).
- Changes in smoking rates and age of initiation (long-term outcome).
- Changes in morbidity and mortality (long-term outcome).

Comparing tobacco-related data among states and between one state and the nation as a whole are common and important ways to evaluate tobacco control programs. Another option is to compare data from different—but relevant—sources. For example, you could make comparisons using indicators from the YTS, the BRFSS tobacco module, PRAMS, and a survey of adult tobacco use. Comparing your data with national data and other states' data will help you to establish realistic objectives for your program and meaningful benchmarks for progress. States can also compare their progress with that of states with a similar investment in tobacco control, or they can contrast their results (outcomes) with the results that could be expected if their program were similar to those of states with a larger investment in tobacco control.

Comparison data are also useful for measuring indicators in anticipation of new or expanding programs. For example, noting a "lack of change" in key indicators over time prior to program implementation helps demonstrate the need for your program and highlights the comparative progress of states with comprehensive tobacco control programs already in place. A lack of change in indicators may continue for several years and is useful as a justification for greater investment in evidence-based, well-funded, and more comprehensive programs. There

are many opportunities for between-state comparisons and trend analysis, which can be highlighted with time-series analyses. The tobacco questions on many of the larger surveillance systems have not changed in several years, so you can make comparisons with other states and over time, using specific indicators. Program managers are encouraged to collaborate with state epidemiologists, BRFSS coordinators, and statisticians to make state and national comparisons an important component of your evaluation.

Common types of evaluation designs

The field of health promotion is under increasing pressure to demonstrate that programs are worthwhile, effective, and efficient. During the last 2 decades, knowledge and understanding about how to evaluate complex programs have increased significantly. The appropriateness of the evaluation design is a primary concern. The evaluation design ought to accommodate the complexity of program activities and meet the needs of diverse stakeholders. As a result, states are often encouraged to use multiple methods to evaluate program efforts. However, "the use of randomized control trials to evaluate health promotion initiatives is, in most cases, inappropriate, misleading, and unnecessarily expensive."[19]

Three general types of evaluation designs are commonly recognized: experimental, quasi-experimental, and observational. Evaluations using experimental designs use random assignment to compare the effect of an intervention on one or more groups with effect on an equivalent group or groups that did not receive the intervention. For example, an evaluation team could select a group of similar schools, then randomly assign some schools to receive a tobacco-use prevention curriculum and other schools to serve as control schools. All schools have the same chance of being selected as an intervention or control school. Because of the "random assignment," you reduce the chances that the control and intervention schools vary in any way that could influence differences in program outcomes. This allows you to attribute change in outcomes to your program. For example, if the students in the intervention schools delayed smoking onset longer than students in the control schools, you could attribute the success to your program.

Sometimes in community settings it is hard, or even unethical, to have a true control group. One solution is to offer the program to the control group after data for the evaluation have been collected. Another option is to use a quasi-experimental design. These designs make comparisons between nonequivalent groups and do not involve random assignment to intervention and control groups. An example would be to assess adults' beliefs about the harmful effects of ETS in two communities, then conduct a media campaign in one of the communities. After the campaign, you would reassess the adults and expect to find a higher percentage of adults believing ETS is harmful in the community that received the media campaign. Critics could argue that other differences between the two communities caused the changes in beliefs, so it is important to document that the intervention and comparison groups are similar on key factors such as population demographics and related current or historical events.

Observational designs are also used in program evaluation. These include, but are not limited to, longitudinal, cross-sectional surveys and case studies. Periodic cross-sectional surveys (e.g., the YTS or BRFSS) can inform your evaluation. Case studies may be particularly appropriate for assessing changes in tobacco control capacity in disparate population groups. Case studies are often applicable when the program is unique, when an existing program is used in a different setting, when you are assessing a unique outcome, or when an environment is especially unpredictable. Case studies can also allow for an exploration of community characteristics and how these may influence program implementation as well as the identification of barriers to and facilitators of change. One resource on case studies is *Using Case Studies To Do Evaluation*, by the California Department of Health Services' Tobacco Control Section (www.dhs.cahwnet.gov/ps/cdic/ccb/TCS/documents/ProgramEvaluation.pdf). This guide can help evaluators determine whether and how to use a case study approach.

Given the widespread visibility of antitobacco messages and overlapping program components, traditional evaluation designs (experimental and quasi-experimental) have proven difficult to implement and hard to maintain. Some tobacco control program outcomes are often detectable only after several years.

Therefore, before choosing an experimental or quasi-experimental design for your evaluation, consider the appropriateness and feasibility of less traditional designs (e.g., simple before-after [pretest-posttest] or posttest-only designs). Depending on your program's objectives and the intended use(s) for the evaluation findings, these designs may be more suitable for measuring progress toward achieving program goals. And these designs often cost less and require less time. Keep in mind, however, that saving time and money should not be the main criterion when selecting an evaluation design. It is important to choose a design that will measure what you need to measure and that will meet both your immediate and long-term needs.

A *goal-based evaluation model* uses predetermined program goals as the standards for evaluation, thus holding the program accountable to prior expectations. In such cases, evaluation planning focuses on the activities, outputs, and short-term, intermediate, and long-term outcomes outlined in a program logic model to direct measurement activities. One advantage of this evaluation model is that the evaluation team has flexibility and can adapt evaluation strategies if notable changes occur in the inputs and activities of the program. In the early stages of your program, progress toward objectives can be measured to document achievement and demonstrate accountability.

The design you select influences the timing of data collection, how you analyze the data, and the types of conclusions you can make from your findings. A collaborative approach to focusing the evaluation provides a practical way to better ensure the appropriateness and utility of your evaluation design.

Purpose

You should articulate the purposes of your evaluation. These may be to improve the program, assess program effectiveness, or demonstrate accountability for resources. The purposes will reflect the stage of development of your program. With a new program, you will probably want to conduct a process evaluation to help improve the program. With a mature program, you will probably want to conduct an outcome evaluation to assess your program's effectiveness and to demonstrate that it is making productive use of resources.

Improving the program

Program evaluation can identify areas in the program that need improvement. For example, a smoking-cessation program may be effective, but it may not be attracting or retaining many participants. By conducting a process evaluation you may discover why. For example, the program may be at an inconvenient location, or participants may not have access to transportation or child care. Cost may be a barrier. As a result, program coordinators may attempt to increase attendance by moving the location of the class, providing free public transportation, working with purchasers and insurers to increase coverage for programs, or switching to a telephone cessation help-line to increase access.

Assessing the program's effectiveness

Program evaluation can measure how effective your program is at progressing toward the desired outcomes. For example, evaluation can assess whether a school-based tobacco prevention program is increasing students' knowledge about the dangers of tobacco, or whether a cessation program is increasing the duration or permanency of participants' attempts to quit smoking. Information about the effectiveness of a program can be used to make decisions about the continuation, refinement, or expansion of the program.

Demonstrating productive use of resources

Program managers are typically accountable to funders and various stakeholders, including government officials and policymakers. Program managers must justify how and where their funds are spent. Evaluation results can be used to demonstrate that a program is functioning as planned, achieving its objectives, worth the cost, or making an important contribution to health.

Defining the users of evaluation results

The evaluation team must also consider who will use the evaluation results. Those users need to be identified and given the opportunity to provide input into the design of the evaluation. Support from the intended users will increase the likelihood that they will use the evaluation results. Users of evaluation findings differ from the larger network of program stakeholders in that the information needs of intended users will determine how you focus the evaluation.

Defining the uses of evaluation results

How your results will be used depends on the purpose and intended users of the evaluation. You need a plan for each piece of information collected. Consider also why you are collecting it and what you are going to do with it. In tobacco control and prevention, evaluation information may be used, for example,

- To identify areas of the program that need improvement.
- To decide how to allocate resources.
- To document the level of success in achieving objectives.
- To assess community needs.
- To mobilize community support.
- To redistribute or expand the locations where the intervention is carried out.
- To improve the content of the program's materials.
- To focus program resources on a specific population.

Evaluation questions

A focused evaluation gathers information for a specific purpose or use. Evaluation questions need to be discussed with and agreed upon by the stakeholders. After you have identified the evaluation users, you must determine what is important to them and design your evaluation questions to meet their needs. Because the questions your evaluation team and stakeholders agree on will affect the methods you use to gather data, you must decide which questions to ask before you choose your methods.

Besides having a specific purpose and use, your evaluation should also reflect the stage of your program's development. For example, you must decide whether you are conducting an outcome evaluation, a process evaluation, or both. Process evaluations and outcome evaluations require different designs and collect different types of data. Think about the stage of your program's development in making these decisions. If you have a well-established program, you may wish to evaluate changes in intermediate or long-term outcomes. However, the evaluation team should determine which outcomes are the most important to evaluate at each stage of program development. Decisions

about the evaluation questions and outcomes you plan to measure should be made by the evaluation team in collaboration with key stakeholders.

✔ Checklist for focusing the evaluation design

- Define the purpose(s) of your evaluation.
- Identify the use(s) of the evaluation results.
- Formulate the questions the evaluation will answer.
- Distinguish evaluation from research questions.
- Review evaluation questions with stakeholders, program managers, and program staff.
- Include process and outcome evaluation.
- Review options for the evaluation design.
- Consider a goal-based evaluation model.
- Make sure that the evaluation design fits the evaluation questions.
- Collect baseline data.
- Plan how to compare your data with those of other states and with national data.
- Consider local or regional comparisons, or both.
- Seek technical expertise or review.
- Document the need for the program.
- Document program resources.
- Note the program's stage of development.
- Explain the program context.
- List and describe program activities.
- State program goals and objectives.
- Prepare a logic model.

Resources

1. CDC Evaluation Working Group
   www.cdc.gov/eval
2. Using Case Studies to do Program Evaluation
   www.dhs.cahwnet.gov/ps/cdic/ccb/TCS/html/Evaluation_Resources.htm
3. Local Program Evaluation Planning Guide
   www.dhs.cahwnet.gov/ps/cdic/ccb/TCS/html/Evaluation_Resources.htm
4. CDC. Strategies for reducing exposure to environmental tobacco smoke, increasing tobacco-use cessation, and reducing initiation in communities and health-care systems. A report on recommendations of the Task Force on Community Preventive Services. *MMWR* 2000;49(No. RR-12).
   www.cdc.gov/tobacco/research_data/environmental/rr4912.pdf
5. CDC. Decline in cigarette consumption following implementation of a comprehensive tobacco prevention and education program—Oregon, 1996–1998. *MMWR* 1999;48:140–3.
   www.cdc.gov/tobacco/research_data/interventions/mm4807.pdf
6. CDC. Declines in lung cancer rates—California, 1988–1997. *MMWR* 2000;49:1066–70.
   www.cdc.gov/tobacco/research_data/health_consequences/ccmm4947.pdf
7. Lois Biener, Jeffrey E Harris, and William Hamilton. Impact of the Massachusetts tobacco control programme: population based trend analysis. *BMJ* 2000; 321:351–4.
   www.bmj.com

The resources listed here include links to some nongovernmental organizations' Web sites. These sites are provided solely as examples. Links do not constitute an endorsement of these organizations' materials or programs by CDC or the federal government. CDC is not responsible for the content of any individual organization's Web pages found at these links.