

Introduction to Program Evaluation

for Public Health Programs

*Evaluating
Appropriate
Antibiotic Use
Programs*

April 2006



DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention



Suggested Citation:

Centers for Disease Control and Prevention. Office of the Director, Office of Strategy and Innovation, & National Center for Infectious Diseases, Division of Bacterial and Mycotic Diseases. Introduction to program evaluation for public health programs: Evaluating appropriate antibiotic use programs. Atlanta, GA: Centers for Disease Control and Prevention, 2006.

About This Manual

This manual is a customized version of the “Introduction to Program Evaluation for Public Health Programs: A Self Study Guide” with specific information and examples for appropriate antibiotic use programs. The original manual was produced by the Centers for Disease Control and Prevention’s Office of Strategy and Innovation and is available online at <http://www.cdc.gov/eval/whatsnew.htm>.

To download a copy of this manual, “Introduction to Program Evaluation for Public Health Programs: Evaluating Appropriate Antibiotic Use Programs,” please visit the appropriate antibiotic use program website at www.cdc.gov/getsmart.

Acknowledgments

From the Office of the Director, Office of Strategy and Innovation:

This manual integrates, in part, the excellent work of the many Centers for Disease Control and Prevention (CDC) programs that have used the CDC’s *Framework for Program Evaluation in Public Health* to develop guidance documents and other materials for their grantees and partners. We thank in particular the Office on Smoking and Health, and the Division of Nutrition and Physical Activity, whose prior work influenced the content of this manual.

We thank the following people from the Evaluation Manual Planning Group for their assistance in coordinating, reviewing, and producing this document. In particular:

NCHSTP, Division of TB Elimination: Maureen Wilce
NCID, Division of Bacterial and Mycotic Diseases: Jennifer Weissman
NCCDPHP, Division of Diabetes Translation: Clay Cooksey
NCEH, Division of Airborne and Respiratory Diseases: Kathy Sunnarborg

We extend special thanks to Daphna Gregg and Antoinette Buchanan for their careful editing and composition work on drafts of the manual, and to the staff of the Office of the Associate Director of Science for their careful review of the manual and assistance with the clearance process.

From the staff of Get Smart: Know When Antibiotics Work:

The sections of this manual on evaluating appropriate antibiotic use programs were written by Jennifer Weissman, with input from the entire Get Smart team, particularly Rich Besser, Alison Patti, Stefanie Anderson, and Erica Haller-Stevenson. The team would like to acknowledge the contributions of Laura Saddler (Oregon), Ben Techagaiciyawanis (Los Angeles County), Elissa

Maas (California), Mary Hallin (Nebraska), and Lisa Bridges in helping make this manual user-friendly and relevant by providing comments on an earlier draft of this manual. We extend thanks to Heidi Brown and Martha Iwamoto for helping pilot test sections of this manual and this approach to evaluation during site visits to Michigan and Missouri during the spring of 2005, and to Mary Eley and the MARR Coalition (Michigan), and Bryan Norman and the ARCH Coalition (Missouri) for their active participation and enthusiasm during on-site evaluation trainings. Finally, we are grateful to Goldie MacDonald and Tom Chapel for introducing us to logic models, encouraging us to write this manual, and providing numerous helpful resources along the way.

Contents

	Page
Executive Summary	
Introduction.....	1
Step 1: Engage Stakeholders.....	13
Step 2: Describe the Program.....	23
Step 3: Focus the Evaluation Design.....	50
Step 4: Gather Credible Evidence.....	66
Step 5: Justify Conclusions.....	84
Step 6: Ensure Use of Evaluation Findings and Share Lessons Learned.....	92
Case Study: Clinic-Based Education for Patients and Providers.....	101
Case Study: Media Campaign.....	110
Glossary.....	118
Program Evaluation Resources.....	127

Executive Summary

This document is a “how to” guide for planning and implementing evaluation activities. The manual is based on CDC’s *Framework for Program Evaluation in Public Health*, and is intended to assist state, local, and community managers and staff of public health programs in planning, designing, implementing, and using the results of comprehensive evaluations in a practical way. The strategy presented in this manual will help assure that evaluations meet the diverse needs of internal and external stakeholders, including assessing and documenting program implementation, outcomes, efficiency, and cost-effectiveness of activities, and taking action based on evaluation results to increase the impact of programs.

Why Evaluate Public Health Programs?

Public health programs have as their ultimate goal preventing or controlling disease, injury, disability, and death. Over time, this task has become more complex as programs themselves have become more complex. Increasingly, public health programs address large problems, the solution to which must engage large numbers of community members and organizations in a vast coalition. More often than not, public health problems—which in the last century might have been solved with a vaccine or change in sanitary systems—involve significant and difficult changes in attitudes and risk/protective behavior of consumers and/or providers.

In addition, the context in which public health programs operate has become more complex. Programs that work well in some settings fail dismally in others because of the fiscal, socioeconomic, demographic, interpersonal, and/or interorganizational setting in which they are planted. At the same time that programs have become more complex, the demands of policymakers and other stakeholders for accountability have increased.

All these changes in the environment in which public health programs operate mean that strong program evaluation is essential now more than ever, but also that there is no one “right” evaluation. Rather, a host of evaluation questions may arise over the life of the program that might reasonably be asked at any point in time. Addressing these questions about program effectiveness means paying attention to documenting and measuring the implementation of the program and its success in achieving intended outcomes, and using such information to be accountable to key stakeholders.

Program Implementation

The task of evaluation encourages us to examine the operations of a program, including which activities take place, who conducts the activities, and who is reached as a result. In addition, evaluation will show how faithfully the program adheres to implementation protocols. Through program evaluation, we can determine whether activities are implemented as planned and identify program strengths, weaknesses, and areas for improvement.

For example, a treatment program may be very effective for those who complete it, but the number of participants may be low. Program evaluation may identify the location of the program or lack of transportation as a barrier to attendance. Armed with this information, program managers can move the class location or meeting times or provide free transportation, thus enhancing the chances the program will actually produce its intended outcomes.

Program Effectiveness

CDC and the federal government have identified goals that public health programs should work toward to prevent or reduce morbidity and mortality. Comprehensive public health programs use multiple strategies to address these goals. Typically, strategies are grouped into program components that might include, for example, community mobilization, policy and regulatory action, strategic use of media and health communication, and funding of frontline programs. Program evaluation includes documenting progress on program goals and the effectiveness of these strategies in producing this progress.

Program Accountability

Program evaluation is a tool with which to demonstrate accountability to the array of stakeholders, who for a given program may include funding sources, policymakers, state and local agencies implementing the program, or community leaders. Depending on the needs of stakeholders, program evaluation findings may demonstrate that the program makes a contribution to reducing morbidity and mortality or relevant risk factors; or that money is being spent appropriately and effectively; or that further funding, increased support, and policy change might lead to even more improved health outcomes. By holding programs accountable in these ways, evaluation helps ensure that the most effective approaches are maintained and that limited resources are spent efficiently.

This manual is based on the CDC's *Framework for Program Evaluation in Public Health*,¹ and integrates insights from Framework-based manuals developed by CDC's Office on Smoking and Health,² and Division of Nutrition and Physical Activity³ for their grantees and state and local partners, and by the Center for the Advancement of Community Based Public Health for community health programs.⁴ This document is organized around the six steps of the CDC Framework:

- Engage Stakeholders
- Describe The Program
- Focus The Evaluation
- Gather Credible Evidence
- Justify Conclusions
- Ensure Use of Evaluation Findings and Share Lessons Learned

Each chapter illustrates the main points using examples inspired by real programs at the federal, state, and local levels. In addition, following each chapter are supplementary materials that apply the main points of the chapter to your specific public health problem or area. These supplementary materials include one or more crosscutting case examples relevant to the specific public health area.

¹ Centers for Disease Control and Prevention. Framework for program evaluation in public health. Atlanta, GA: MMWR 1999;48(NoRR-11):1-40.

² US Department of Health and Human Services. Introduction to program evaluation for comprehensive tobacco control programs. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, Office on Smoking and Health, November 2001.

³ US Department of Health and Human Services. Physical activity evaluation handbook. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2002.

⁴ Center for Advancement of Community Based Public Health. An evaluation framework for community health programs. Durham, NC: Center for Advancement of Community Based Public Health, June 2000.

Introduction

What is Program Evaluation?

Most program managers assess the value and impact of their work all the time when they ask questions, consult partners, make assessments, and obtain feedback. They then use the information collected to improve the program. Indeed, such informal assessments fit nicely into a broad definition of evaluation as the “*examination of the worth, merit, or significance of an object.*”⁵ And throughout this manual, the term “program” will be defined as “*any set of organized activities supported by a set of resources to achieve a specific and intended result.*” This definition is intentionally broad so that almost any organized public health action can be seen as able to benefit from program evaluation:

- Direct service interventions (e.g., a program that offers free breakfasts to improve nutrition for grade school children)
- Community mobilization efforts (e.g., an effort to organize a boycott of California grapes to improve the economic well-being of farm workers)
- Research initiatives (e.g., an effort to find out whether disparities in health outcomes based on race can be reduced)
- Advocacy work (e.g., a campaign to influence the state legislature to pass legislation regarding tobacco control)
- Training programs (e.g., a job training program to reduce unemployment in urban neighborhoods)

What makes true program evaluation different from the sort of informal assessment that any smart and dedicated manager is doing all the time? Mainly, it is that evaluation is conducted according to a set of guidelines (protocols) that are systematic, consistent, and comprehensive to assure the accuracy of the results. For purposes of this manual, we will define program evaluation as “*the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future program development.*”⁶ Program evaluation does not occur in a vacuum; rather, it is influenced by real-world constraints. Evaluation should be practical and feasible and must be conducted within the confines of resources, time, and political context. Moreover, it should serve a useful purpose, be conducted in an ethical manner, and produce accurate findings. Evaluation findings should be used both to make decisions about program implementation and to improve program effectiveness.

As you will see, many different questions can be part of a program evaluation, depending on how long the program has been in existence, who is asking the question, and why the information is needed. In general, evaluation questions fall into one of these groups:

- **Implementation:** Were your program’s activities put into place as originally intended?

⁵ Scriven M. Minimalist theory of evaluation: The least theory that practice requires. *American Journal of Evaluation* 1998;19:57-70.

⁶ Patton MQ. *Utilization-focused evaluation: The new century text*. 3rd ed. Thousand Oaks, CA: Sage, 1997.

- **Effectiveness:** Is your program achieving the goals and objectives it was intended to accomplish?
- **Efficiency:** Are your program's activities being produced with appropriate use of resources such as budget and staff time?
- **Cost-Effectiveness:** Does the value or benefit of achieving your program's goals and objectives exceed the cost of producing them?
- **Attribution:** Can progress on goals and objectives be shown to be related to your program, as opposed to other things that are going on at the same time?

All of these are appropriate evaluation questions and might be asked with the intention of documenting program progress, demonstrating accountability to funders and policymakers, or identifying ways to make the program better.

Evaluation Supplements Other Types of Reflection and Data Collection

Evaluation is one of several ways in which the staff of a program might answer the question “How are we doing?” In most large organizations, that question might be posed at budgeting time, during strategic planning, in constructing performance measures, or even in establishing the marketing “brand” for the organization. And the question might be answered using approaches that could be characterized as “surveillance,” “research,” or “program evaluation.” It is important that organizations see these processes as related and do their best to integrate the insights from them. Here’s how:

Planning

Planning asks, “What *are* we doing and what *should* we do to achieve our goals?” Program evaluation, by providing information on progress toward organizational goals and identifying which parts of the program are working well and/or poorly, sets up the discussion of what can be changed to help the program better meet its intended goals and objectives.

Performance Measurement

Increasingly, public health programs are called to be accountable to funders, legislators, and the general public. Many programs do this by creating, monitoring, and reporting results for a small set of markers and milestones of program progress. Such “performance measures” are a type of evaluation—answering the question “*How* are we doing?” More importantly, when performance measures show significant or sudden changes in program performance, program evaluation efforts can be directed to the troubled areas to determine “*Why* are we doing poorly or well?”

Budgeting

Linking program performance to program budget is the final step in accountability. Called “activity-based budgeting” or “performance budgeting,” it requires an understanding of program components and the links between activities and intended outcomes. The early steps in the program evaluation approach (such as logic modeling) clarify these relationships, making the link between budget and performance easier and more apparent.

Surveillance and Program Evaluation

While the terms *surveillance* and *evaluation* are often used together, each makes a distinctive contribution to a program, and it is important to clarify their different purposes. *Surveillance* is the continuous monitoring or routine data collection on various factors (e.g., behaviors, attitudes, deaths) over a regular interval of time. Surveillance systems have existing resources and infrastructure. Data gathered by surveillance systems are invaluable for performance measurement and program evaluation, especially of longer term and population-based outcomes. In addition, these data serve an important function in program planning and “formative” evaluation by identifying key burden and risk factors—the descriptive and analytic epidemiology of the public health problem. There are limits to how useful surveillance data can be for evaluators. For example, some surveillance systems such as the Behavioral Risk Factor Surveillance System (BRFSS), Youth Tobacco Survey (YTS), and Youth Risk Behavior Survey (YRBS) can measure changes in large populations, but have insufficient sample sizes to detect changes in outcomes for more targeted programs or interventions. Also, these surveillance systems may have limited flexibility when it comes to adding questions that a particular program evaluation might like to have answered.

In the best of all worlds, surveillance and evaluation are companion processes that can be conducted simultaneously. *Evaluation* may supplement surveillance data by providing tailored information to answer specific questions about a program. Data collection that flows from the specific questions that are the focus of the evaluation is more flexible than surveillance and may allow program areas to be assessed in greater depth. For example, a state may supplement surveillance information with detailed surveys to evaluate how well a program was implemented and the impact of a program on participants’ knowledge, attitudes, and behavior. They can also use qualitative methods (e.g., focus groups, feedback from program participants from semi-structured or open-ended interviews) to gain insight into the strengths and weaknesses of a particular program activity.

Research and Program Evaluation

Both research and program evaluation make important contributions to the body of knowledge, but fundamental differences in the purpose of research and the purpose of evaluation mean that good program evaluation need not always follow an academic research model. Even though some of these differences have tended to break down as research tends toward increasingly participatory models⁷ and some evaluations aspire to make statements about attribution, “pure” research and evaluation serve somewhat different purposes (“Distinguishing Principles of Research and Evaluation” table), nicely summarized in the adage “Research seeks to prove; evaluation seeks to improve.” Academic research focuses primarily on testing hypotheses; a key purpose of program evaluation is to improve practice. Research is generally thought of as requiring a controlled environment or control groups. In field settings directed at prevention and control of a public health problem, this is seldom realistic. The last three attributes in the table are especially worth noting. Unlike pure academic research models, program evaluation

⁷ Green LW, George MA, Daniel M, Frankish CJ, Herbert CP, Bowie WR, et al. Study of participatory research in health promotion: Review and recommendations for the development of participatory research in health promotion in Canada. Ottawa, Canada: Royal Society of Canada, 1995.

acknowledges and incorporates differences in values and perspectives from the start, may address many questions besides attribution, and tends to produce results for varied audiences.

Distinguishing Principles of Research and Evaluation

Concept	Research Principles	Program Evaluation Principles
Planning	Scientific method <ul style="list-style-type: none"> State hypothesis. Collect data. Analyze data. Draw conclusions. 	Framework for program evaluation <ul style="list-style-type: none"> Engage stakeholders. Describe the program. Focus the evaluation design. Gather credible evidence. Justify conclusions. Ensure use and share lessons learned.
Decision Making	Investigator-controlled <ul style="list-style-type: none"> Authoritative. 	Stakeholder-controlled <ul style="list-style-type: none"> Collaborative.
Standards	Validity <ul style="list-style-type: none"> Internal (accuracy, precision). External (generalizability). 	Repeatability program evaluation standards <ul style="list-style-type: none"> Utility. Feasibility. Propriety. Accuracy.
Questions	Facts <ul style="list-style-type: none"> Descriptions. Associations. Effects. 	Values <ul style="list-style-type: none"> Merit (i.e., quality). Worth (i.e., value). Significance (i.e., importance).
Design	Isolate changes and control circumstances <ul style="list-style-type: none"> Narrow experimental influences. Ensure stability over time. Minimize context dependence. Treat contextual factors as confounding (e.g., randomization, adjustment, statistical control). Understand that comparison groups are a necessity. 	Incorporate changes and account for circumstances <ul style="list-style-type: none"> Expand to see all domains of influence. Encourage flexibility and improvement. Maximize context sensitivity. Treat contextual factors as essential information (e.g., system diagrams, logic models, hierarchical or ecological modeling). Understand that comparison groups are optional (and sometimes harmful).
Data Collection	Sources <ul style="list-style-type: none"> Limited number (accuracy preferred). Sampling strategies are critical. Concern for protecting human subjects. Indicators/Measures <ul style="list-style-type: none"> Quantitative. Qualitative. 	Sources <ul style="list-style-type: none"> Multiple (triangulation preferred). Sampling strategies are critical. Concern for protecting human subjects, organizations, and communities. Indicators/Measures <ul style="list-style-type: none"> Mixed methods (qualitative, quantitative, and integrated).
Analysis & Synthesis	Timing <ul style="list-style-type: none"> One-time (at the end). Scope <ul style="list-style-type: none"> Focus on specific variables. 	Timing <ul style="list-style-type: none"> Ongoing (formative and summative). Scope <ul style="list-style-type: none"> Integrate all data.
Judgments	Implicit <ul style="list-style-type: none"> Attempt to remain value-free. 	Explicit <ul style="list-style-type: none"> Examine agreement on values. State precisely whose values are used.
Conclusions	Attribution <ul style="list-style-type: none"> Establish time sequence. Demonstrate plausible mechanisms. Control for confounding. Replicate findings. 	Attribution and contribution <ul style="list-style-type: none"> Establish time sequence. Demonstrate plausible mechanisms. Account for alternative explanations. Show similar effects in similar contexts.
Uses	Disseminate to interested audiences <ul style="list-style-type: none"> Content and format varies to maximize comprehension. 	Feedback to stakeholders <ul style="list-style-type: none"> Focus on intended uses by intended users. Build capacity. Disseminate to interested audiences <ul style="list-style-type: none"> Content and format varies to maximize comprehension. Emphasis on full disclosure. Requirement for balanced assessment.

Why Evaluate Public Health Programs?

Some Reasons to Evaluate Public Health Programs

- To monitor progress toward the program's goals.
- To determine whether program components are producing the desired progress on outcomes.
- To permit comparisons among groups, particularly among populations with disproportionately high risk factors and adverse health outcomes.
- To justify the need for further funding and support.
- To find opportunities for continuous quality improvement.
- To ensure that effective programs are maintained and resources are not wasted on ineffective programs.

Program staff may be *pushed* to do evaluation by external mandates from funders, authorizers, or others, or they may be *pulled* to do evaluation by an internal need to determine how the program is performing and what can be improved. While push *or* pull can motivate a program to conduct good evaluations, program evaluation efforts are more likely to be sustained when staff see the results as useful information that can help *them* do their jobs better.

Data gathered during evaluation enable managers and staff to create the best possible programs, to learn from mistakes, to make modifications as needed, to monitor progress toward program goals, and to judge the success of the program in achieving its short-term, intermediate, and long-term outcomes. Most public health programs aim to change behavior in one or more target groups and to create an environment that reinforces sustained adoption of these changes, with the intention that changes in environments and behaviors will prevent and control diseases and injuries. Through evaluation, you can track these changes and, with careful evaluation designs, assess the effectiveness and impact of a particular program, intervention, or strategy in producing these changes.

Recognizing the importance of evaluation in public health practice and the need for appropriate methods, the World Health Organization (WHO) established the Working Group on Health Promotion Evaluation. The Working Group prepared a set of conclusions and related recommendations to guide policymakers and practitioners.⁸ Recommendations immediately relevant to the evaluation of comprehensive public health programs include:

- Encourage the adoption of participatory approaches to evaluation that provide meaningful opportunities for involvement by all of those with a direct interest in initiatives (programs, policies, and other organized activities).
- Require that a portion of total financial resources for a health promotion initiative be allocated to evaluation—they recommend 10%.
- Ensure that a mixture of process and outcome information is used to evaluate all health promotion initiatives.
- Support the use of multiple methods to evaluate health promotion initiatives.

⁸ WHO European Working Group on Health Promotion Evaluation. Health promotion evaluation: Recommendations to policy-makers: Report of the WHO European working group on health promotion evaluation. Copenhagen, Denmark: World Health Organization, Regional Office for Europe, 1998.

- Support further research into the development of appropriate approaches to evaluating health promotion initiatives.
- Support the establishment of a training and education infrastructure to develop expertise in the evaluation of health promotion initiatives.
- Create and support opportunities for sharing information on evaluation methods used in health promotion through conferences, workshops, networks, and other means.

CDC’s Framework for Program Evaluation in Public Health

Program evaluation is 1 of 10 essential public health services⁹ and a critical organizational practice in public health.¹⁰ Until recently, however, there has been little agreement among public health officials on the principles and procedures for conducting such studies. In 1999, CDC published the *Framework for Program Evaluation in Public Health* and some related recommendations.¹¹ The Framework, as depicted in Figure 1.1, defined six steps and four sets of standards for conducting good evaluations of public health programs.



Figure 1.1
Evaluation Framework

The underlying logic of the Framework is that good evaluation does not merely gather accurate evidence and draw valid conclusions, but produces results that are *used* to make a difference. To maximize the chances evaluation results will be used, you need to create a “market” before you create the “product”—the evaluation. You determine the market by focusing your evaluations on questions that are most salient, relevant, and important. And you ensure the best evaluation focus by understanding where the questions fit into the full landscape of your program description, and especially by ensuring that you have identified and engaged stakeholders who care about these questions and want to take action on the results.

The steps in the CDC Framework are informed by a set of standards for evaluation.¹² These standards do not constitute a *way* to do evaluation; rather, they serve to *guide* your choice from among the many options available at each step in the Framework. The 30 standards cluster into four groups:

- **Utility:** Who needs the evaluation results? Will the evaluation provide relevant information in a timely manner for them?
- **Feasibility:** Are the planned evaluation activities realistic given the time, resources, and expertise at hand?

⁹ Public Health Functions Steering Committee. Public health in America. Fall 1994. Available at <<http://www.health.gov/phfunctions/public.htm>>. January 1, 2000.

¹⁰ Dyal WW. Ten organizational practices of public health: A historical perspective. *American Journal of Preventive Medicine* 1995;11(6)Suppl 2:6-8.

¹¹ Centers for Disease Control and Prevention. op cit.

¹² Joint Committee on Standards for Educational Evaluation. *The program evaluation standards: How to assess evaluations of educational programs*. 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.

- **Propriety:** Does the evaluation protect the rights of individuals and protect the welfare of those involved? Does it engage those most directly affected by the program and changes in the program, such as participants or the surrounding community?
- **Accuracy:** Will the evaluation produce findings that are valid and reliable, given the needs of those who will use the results?

Sometimes the standards broaden your exploration of choices; as often, they help reduce the options at each step to a manageable number. For example, in the step “Engaging Stakeholders,” the standards can help you think broadly about who constitutes a stakeholder for your program, but simultaneously can reduce the potential list to a manageable number by posing the following questions based on the standards: (**Utility**) Who will use these results? (**Feasibility**) How much time and effort can be devoted to stakeholder engagement? (**Propriety**) To be ethical, which stakeholders need to be consulted, for example, those served by the program or the community in which it operates? (**Accuracy**) How broadly do you need to engage stakeholders to paint an accurate picture of this program?

Similarly, there are unlimited ways to “gather credible evidence.” Asking these same kinds of questions as you approach evidence gathering will help identify ones that will be most useful, feasible, proper, and accurate for *this* evaluation at *this* time. Thus, the CDC Framework approach supports the fundamental insight that there is no such thing as *the* right program evaluation. Rather, over the life of a program, any number of evaluations may be appropriate, depending on the situation.

How to Select a Lead Evaluator and Establish an Evaluation Team

Good evaluation requires a combination of skills that are rarely found in a single person. An evaluation team that includes internal program staff, external stakeholders, and possibly consultants or contractors with evaluation expertise is the preferred approach. An initial step in the formation of a team is to decide who will be responsible for planning and implementing evaluation activities. At least one program staff person should be selected as the lead evaluator to coordinate program efforts. This person should be responsible for evaluation activities, including planning and budgeting for evaluation, developing program objectives, addressing data collection needs, reporting findings, and working with consultants. The lead evaluator is ultimately responsible for engaging stakeholders, consultants, and other collaborators who bring the skills and interests needed to plan and conduct the evaluation.

Although this staff person should have the skills necessary to competently coordinate evaluation activities, he or she can choose to look elsewhere for technical expertise to design and implement specific tasks. However, developing in-house evaluation expertise and capacity is a beneficial goal for most public health organizations.

Of the characteristics of a good evaluator listed in the accompanying text box, the evaluator’s ability to work with a diverse group of stakeholders warrants highlighting. The lead evaluator should be willing and able to draw out and reconcile differences in values and standards of different stakeholders and to work with knowledgeable stakeholder representatives in designing and conducting the evaluation.

Characteristics of a Good Evaluator

- Has experience in the type of evaluation needed.
- Is comfortable with qualitative and quantitative data sources and analysis.
- Is able to work with a wide variety of stakeholders, including representatives of target populations.
- Can develop innovative approaches to evaluation while considering the realities affecting a program (e.g., a small budget).
- Incorporates evaluation into all program activities.
- Understands both the potential benefits and risks of evaluation.
- Educates program personnel about designing and conducting the evaluation.
- Will give staff the full findings (i.e., will not gloss over or fail to report certain findings for any reason).
- Has strong coordination and organization skills.
- Explains material clearly and patiently.
- Respects all levels of personnel.
- Communicates well with key personnel.
- Exhibits cultural competence.
- Delivers reports and protocols on time.

Additional evaluation expertise sometimes can be found in programs within the health department, through external partners (e.g., universities, organizations, companies), from peer programs in other states and localities, and through technical assistance offered by CDC.¹³

You can also use outside consultants as volunteers, advisory panel members, or contractors. External consultants can provide high levels of evaluation expertise from an objective point of view. Important factors to consider when selecting consultants are their level of professional training, experience, and ability to meet your needs. Overall, it is important to find a consultant whose approach to evaluation, background, and training best fit your program's evaluation needs and goals. Be sure to check all references carefully before you enter into a contract with any consultant.

To generate discussion around evaluation planning and implementation, several states have formed evaluation advisory panels. Advisory panels typically generate input from local, regional, or national experts otherwise difficult to access. Such an advisory panel will lend additional credibility to your efforts and prove useful in cultivating widespread support for evaluation activities.

The evaluation team members should clearly define their respective roles. Informal consensus may be enough; others prefer a written agreement that describes who will conduct the evaluation and assigns specific roles and responsibilities to individual team members. Either way, the team must clarify and reach consensus on the

- Purpose of the evaluation

¹³ CDC's Prevention Research Centers (PRC) program is an additional resource. The PRC program is a national network of 24 academic research centers committed to prevention research and the ability to translate that research into programs and policies. The centers work with state health departments and members of their communities to develop and evaluate state and local interventions that address the leading causes of death and disability in the nation. Additional information on the PRCs is available at www.cdc.gov/prc/index.htm.

- Potential users of the evaluation findings and plans for dissemination
- Evaluation approach
- Resources available
- Protection for human subjects.

The agreement should also include a timeline and a budget for the evaluation.

Organization of This Manual

This manual is organized by the six steps of the CDC Framework. Each chapter will introduce the key questions to be answered in that step, approaches to answering those questions, and how the four evaluation standards might influence your approach. The main points are illustrated with one or more public health examples that are composites inspired by actual work being done by CDC, states, and localities.¹⁴ Some examples that will be referred to throughout this manual:

Affordable Home Ownership Program

The program aims to provide affordable home ownership to low-income families by identifying and linking funders/sponsors, construction volunteers, and eligible families. Together, they build a house over a multi-week period. At the end of the construction period, the home is sold to the family using a no-interest loan.

Childhood Lead Poisoning Prevention (CLPP)

Lead poisoning is the most widespread environmental hazard facing young children, especially in older inner-city areas. Even at low levels, elevated blood lead levels (EBLL) have been associated with reduced intelligence, medical problems, and developmental problems. The main sources of lead poisoning in children are paint and dust in older homes with lead-based paint. Public health programs address the problem through a combination of primary and secondary prevention efforts. A typical secondary prevention program at the local level does outreach and screening of high-risk children, identifying those with EBLL, assessing their environments for sources of lead, and case managing both their medical treatment and environmental corrections. However, these programs must rely on others to accomplish the actual medical treatment and the reduction of lead in the home environment.

Provider Education in Immunization

A common initiative of state immunization programs is comprehensive provider education programs to train and motivate private providers to provide more immunizations. A typical program includes a newsletter distributed three times per year to update private providers on new developments and changes in policy, and provide a brief education on various immunization topics; immunization trainings held around the state conducted by teams of state program staff and physician educators on general immunization topics and the immunization registry; a Provider Tool Kit on how to increase immunization rates in their practice; training of nursing staff in local health departments who then conduct immunization presentations in individual

¹⁴ These cases are composites of multiple CDC and state and local efforts that have been simplified and modified to better illustrate teaching points. While inspired by real CDC and community programs, they are not intended to reflect the current operation of these programs.

private provider clinics; and presentations on immunization topics by physician peer educators at physician grand rounds and state conferences.

At the conclusion of each chapter are three resources:

- Worksheets to help you apply the teaching points
- Customized information developed by your CDC program on applying the main points of the chapter to your particular public health program
- One or more detailed “worked cases” developed by your CDC program to illustrate how to apply the main points of the chapter to your public health program

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Introduction

The Institute of Medicine has identified antibiotic resistance as one of the key microbial threats to health in the United States. It has listed decreasing the inappropriate use of antimicrobials as a primary solution to address this threat.¹⁵ During the latter half of the 1990s, CDC, many state and local health departments, and other organizations responded to increases in antibiotic resistance and inappropriate prescribing by designing and implementing interventions to promote appropriate antibiotic prescribing in the community. These efforts appear to have contributed to recent decreases in prescribing rates. It is important to evaluate these programs in order to learn which components are most effective and to decide how to use limited resources to continue these efforts.

Antibiotic Resistance and Upper Respiratory Infections

Widespread use of antibiotics has resulted in the development of antibiotic resistance. When antibiotics are used, selective pressure favors the growth of organisms that are resistant to the drug's action. Today, virtually all bacterial pathogens that cause infections of public health importance in the United States and throughout the world are becoming resistant.

Numerous studies have documented the association between recent antibiotic use and carriage of nonsusceptible bacteria. Children who are colonized with *Streptococcus pneumoniae* and have recently received an antibiotic are two to seven times more likely to be colonized with a drug-resistant strain than are children without recent antibiotic use.¹⁶ Research in Alaska showed that increased antibiotic use was strongly associated with an increased likelihood that a person would carry penicillin-resistant bacteria.¹⁷ Furthermore, each additional course of antibiotics was associated with a 20% increase in the odds of carrying an antibiotic-nonsusceptible isolate versus an antibiotic-susceptible isolate. Antibiotic use affects the community as well as the individual.¹⁸ When high levels of antibiotics are used in a community, resistant strains of *Streptococcus pneumoniae* are likely to be circulating. As a result, community members are twice as likely to develop a resistant infection as are people in communities with lower levels of antibiotic use.¹⁹

¹⁵ Institute of Medicine. Microbial threats to health: emergence, detection, and response. March 2003.

¹⁶ Dowell S, Schwartz B. Resistant pneumococci: Protecting patients through judicious use of antibiotics. *American Family Physician* 1997;55:1647-1654.

¹⁷ Hennessy TW, Petersen KM, Bruden D, et al. Changes in antibiotic-prescribing practices and carriage of penicillin-resistant *Streptococcus pneumoniae*: A controlled intervention trial in rural Alaska. *Clinical Infectious Diseases* 2002;34:1543-50.

¹⁸ Emmer CL, Besser RE: Combating antimicrobial resistance: intervention programs to promote appropriate antibiotic use. *Infectious Medicine* 2002;19(4):160-173.

¹⁹ Schwartz G, Kolczak M, Whitney C, et al: US counties with higher rates of antibiotic use have significantly higher proportions of beta lactam and macrolide nonsusceptible *S pneumoniae* antimicrobial resistance. In: Abstracts of the 38th Interscience Conference on Antimicrobial Agents and Chemotherapy. Abstract C-29, San Diego, September 24-27, 1998.

Upper respiratory infections account for three quarters of all antibiotics prescribed by office-based physicians.²⁰ If these antibiotics were being used appropriately, the current increases in resistance rates could be seen as the inevitable consequence of proper treatment. However, this is not the case. Based on data from the 1992 National Ambulatory Medical Care Survey, a population-based survey of prescribing in physicians' offices in the United States, CDC estimated that antibiotic prescribing for upper respiratory infections could be reduced by more than 40%.¹⁸ CDC now estimates that tens of millions of courses of antibiotics are prescribed inappropriately each year for upper respiratory infections.

Get Smart: Know When Antibiotics Work, CDC's national appropriate antibiotic use campaign, targets the five respiratory conditions that in 1992 accounted for more than 75% of all office-based prescribing for all ages combined: otitis media, sinusitis, pharyngitis, bronchitis, and the common cold.²⁰ CDC's appropriate antibiotic use efforts aim to reduce the spread of antibiotic resistance by:

- 1) promoting adherence to appropriate prescribing guidelines among providers,
- 2) decreasing demand for antibiotics for viral upper respiratory infections among healthy adults and parents of young children, and
- 3) increasing adherence to prescribed antibiotics for upper respiratory infections.

Why Evaluate Appropriate Antibiotic Use Programs?

While many states, local health departments, and other groups have implemented appropriate antibiotic use programs aimed at reducing unnecessary antibiotic use and slowing the spread of antibiotic resistance, the vast majority of these programs have not been rigorously evaluated. Since 1998, several controlled trials promoting appropriate antibiotic use for outpatient respiratory tract infections in a variety of settings in the United States have been published.^{17,21,22,23,24} Data from these trials have shown that educational interventions targeting both providers and patients can result in reduced prescribing for respiratory tract infections.

²⁰ McCaig L, Hughes J. Trends in antimicrobial drug prescribing among office-based physicians in the United States. *JAMA* 1995;273:214-219.

²¹ Belongia EA, Sullivan BJ, Chyou PH, et al: A community intervention trial to promote judicious antibiotic use and reduce penicillin-resistant *Streptococcus pneumoniae* carriage in children. *Pediatrics* 2001;108:575-583.

²² Perz JF, Craig AS, Coffey CS, et al. Changes in antibiotic prescribing for children after a community-wide campaign. *JAMA* 2002;287:3103-9.

²³ Gonzales R, Steiner JF, Lum A, et al: Decreasing antibiotic use in ambulatory practice: impact of a multidimensional intervention on the treatment of uncomplicated acute bronchitis in adults. *JAMA* 1999;281:1512-9.

²⁴ Finkelstein JA, Davis RL, Dowell SF, et al: Reducing antibiotic use in children: a randomized trial in 12 practices. *Pediatrics* 2001;108:1-7.

Step 1: Engage Stakeholders

The first step in the CDC Framework approach to program evaluation is to engage the stakeholders. Stakeholders are people or organizations that are invested in the program, are interested in the results of the evaluation, and/or have a stake in what will be done with the results of the evaluation. Representing their needs and interests throughout the process is fundamental to good program evaluation.

Typical Stakeholders in Public Health

Key stakeholders for evaluations of public health programs fall into three major groups:

- Those involved in *program operations*: Management, program staff, partners, funding agencies, and coalition members.
- Those *served or affected* by the program: Patients or clients, advocacy groups, community members, and elected officials.
- Those who are intended *users* of the evaluation findings: Persons in a position to make decisions about the program, such as partners, funding agencies, coalition members, and the general public or taxpayers.

Clearly, these categories are not mutually exclusive; in particular, the primary users of evaluation findings are often members of the other two groups, i.e., the program management or an advocacy organization or coalition. While you may think you know your stakeholders well, these categories help you to think broadly and inclusively in identifying stakeholders.

Potential Stakeholders in Public Health Programs

- Program managers and staff.
- Local, state, and regional coalitions interested in the public health issue.
- Local grantees of your funds.
- Local and national advocacy partners.
- Other funding agencies, such as national and state governments.
- State or local health departments and health commissioners.
- State education agencies, schools, and other educational groups.
- Universities and educational institutions.
- Local government, state legislators, and state governors.
- Privately owned businesses and business associations.
- Health care systems and the medical community.
- Religious organizations.
- Community organizations.
- Private citizens.
- Program critics.
- Representatives of populations disproportionately affected by the problem.
- Law enforcement representatives.

Why Stakeholders are Important to an Evaluation

Stakeholders can help (or hinder) an evaluation *before* it is conducted, *while* it is being conducted, and *after* the results are collected and ready for use. Because so many public health efforts are complex and because public health agencies may be several layers removed from frontline implementation, stakeholders take on particular importance in ensuring that the right evaluation questions are identified and that evaluation results will be used to make a difference. Stakeholders are much more likely to support the evaluation and act on the results and recommendations if they are involved in the evaluation process. Conversely, without stakeholder support, your evaluation may be ignored, criticized, resisted, or even sabotaged.

In reviewing the long list of stakeholders that might be generated in the three generic categories, use of some or all of the evaluation standards will help identify those who matter most.

Use of results will be enhanced if you give priority to those stakeholders who

- Can increase the *credibility* of your efforts or your evaluation
- Are responsible for day-to-day *implementation* of the activities that are part of the program
- Will *advocate* for or *authorize changes* to the program that the evaluation may recommend
- Will *fund* or *authorize the continuation or expansion* of the program.

In addition, to be proper/ethical and accurate, you need to include those who participate in the program and are affected by the program or its evaluation.

The worksheets at the end of this chapter are intended to help you identify key stakeholders. For example, in using the worksheets with the Childhood Lead Poisoning Prevention (CLPP) program, we identified the stakeholders in the sample worksheet 1A (see Table 1.1). Note that some stakeholders appear in more than one column; these are not exclusive classes of stakeholders so much as four ways of thinking about stakeholders to ensure we were thinking as broadly as possible. Second, note that not all categories have the same number of stakeholders. Indeed, for a simple project, there may be very few stakeholders and some categories may have none at all. The sample worksheet 1B (see Table 1.2) helped us identify the perspectives and needs of these key stakeholders and the implications for designing and implementing our evaluation. Note in the CLPP example that while all stakeholders may applaud our efforts to reduce EBLL in children, several stakeholders put priority on outcomes that might or might not agree with our priorities. For example, private physicians are most interested in “yield” of their screening efforts, while Congress cares about cost-effectiveness. Note that advocacy groups, in addition to specific outcomes that may be priorities for them, also have some preferences related to data collection—expressing a preference for methods other than surveys. All of these insights are helpful at the start of an evaluation to ensure that the evaluation goes smoothly and the results are used.

Table 1.1
CLPP Example: Identifying Stakeholders

Who are the key stakeholders we need to:			
Increase credibility of our efforts	Implement the interventions that are central to this effort	Advocate for changes to institutionalize this effort	Fund/authorize continuation or expansion of this effort
Physician associations Community associations	State and local health departments Housing authorities	Advocacy groups Maternal and child health groups Physician associations Community associations	Legislators and policymakers at federal and state levels CDC Private industry Court system

Table 1.2
CLPP Example: What Matters to Stakeholders

Stakeholders		What component of intervention/outcome matters most to them
1	Physician associations	Sufficient “yield” of EBLL children to make their screening efforts “worth their time.” Clear referral mechanisms that are easy and work.
2	Community associations	Cleaning up housing in their neighborhood. Support for families with EBLL children.
3	Housing authorities	No additional monetary and time burden for toxic clean-ups.
4	State and local health departments	Efforts lead to improved health outcome for EBLL children.
5	Advocacy groups	EBLL is seen as a housing problem and not a “failure” or example of bad child-rearing by poor families. No survey data collection with families.
6	Congress and policymakers	Efforts lead to improved health outcomes. “Cost-effectiveness” of the effort.

What to Ask Stakeholders

Throughout the evaluation planning process, you will be asking some or all stakeholders the following questions:

- Who do you represent and why are you interested in this program?
- What is important about this program to you?
- What would you like this program to accomplish?
- How much progress would you expect this program to have made at this time?
- What do you see as the critical evaluation questions at this time?

- How will you use the results of this evaluation?
- What resources (i.e., time, funds, evaluation expertise, access to respondents, and access to policymakers) might you contribute to this evaluation effort?

The Role of Stakeholders in an Evaluation

Stakeholder perspectives may influence every step of the CDC Framework. Obviously, stakeholder input in “describing the program” ensures a clear and consensual understanding of the program’s activities and outcomes. This is an important backdrop for even more valuable stakeholder input in “focusing the evaluation design” to ensure that the key questions of most importance will be included. Stakeholders may also have insights or preferences on the most effective and appropriate ways to collect data from target respondents. In “justifying conclusions,” the perspectives and values that stakeholders bring to the project are explicitly acknowledged and honored in making judgments about evidence gathered. Finally, the considerable time and effort spent in engaging and building consensus among stakeholders pays off in the last step, “ensuring use,” because stakeholder engagement has created a market for the evaluation results. Stakeholders can be involved in the evaluation at various levels. For example, you may want to include coalition members on an evaluation team and engage them in developing questions, data collection, and analysis. Or consider ways to assess your partners’ needs and interests in the evaluation, and develop means of keeping them informed of its progress and integrating their ideas into evaluation activities. Again, stakeholders are more likely to support the evaluation and act on results and recommendations if they are involved in the evaluation process.

In addition, it can be beneficial to engage your program’s critics in the evaluation. In some cases, these critics can help identify issues around your program strategies and evaluation information that could be attacked or discredited, thus helping you strengthen the evaluation process. This information might also help you and others understand the opposition’s rationale and could help you engage potential agents of change within the opposition. However, use caution: It is important to understand the motives of the opposition before engaging them in any meaningful way.

This emphasis on engaging stakeholders mirrors the increasing prominence in the research community of participatory models or “action” research. A participatory approach combines systematic inquiry with the collaboration of diverse stakeholders to meet specific needs and to contend with broad issues of equity and justice. As noted earlier, *The Study of Participatory Research in Health Promotion*, commissioned by the Royal Society of Canada, has published a set of guidelines for use by evaluators and funding agencies in assessing projects that aspire to be participatory.²⁵ The guidelines emphasize that traditional ways of conducting health research in populations must adapt to meet the educational, capacity-building, and policy expectations of more participatory approaches if the results of the research are to make a difference.

²⁵ Green LW, George MA, Daniel M, Frankish CJ, Herbert CP, Bowie WR, et al. op cit.

Standards for Step 1: Engage Stakeholders

Standard	Questions
Utility	<ul style="list-style-type: none">• Who will use these results?
Feasibility	<ul style="list-style-type: none">• How much time and effort can be devoted to stakeholder engagement?
Propriety	<ul style="list-style-type: none">• Which stakeholders need to be consulted to conduct an ethical evaluation, for example, to ensure we will identify negative as well as positive aspects of the program?
Accuracy	<ul style="list-style-type: none">• How broadly do we need to engage stakeholders to paint an accurate picture of this program?

Checklist for Engaging Stakeholders

- Identify stakeholders, using the three broad categories discussed: those affected, those involved in operations, and those who will use the evaluation results.
- Review the initial list of stakeholders to identify key stakeholders needed to improve credibility, implementation, advocacy, or funding/authorization decisions.
- Engage individual stakeholders and/or representatives of stakeholder organizations.
- Create a plan for stakeholder involvement and identify areas for stakeholder input.
- Target selected stakeholders for regular participation in key steps, including writing the program description, suggesting evaluation questions, choosing evaluation questions, and disseminating evaluation results.

**Worksheet 1A
Identifying Key Stakeholders**

Category		Stakeholders
1	Who is affected by the program?	
2	Who is involved in program operations?	
3	Who will use evaluation results?	

Which of these are key stakeholders we need to engage to:			
Increase <u>credibility</u> of our evaluation	<u>Implement</u> the interventions that are central to this evaluation	<u>Advocate</u> for changes to institutionalize the evaluation findings	<u>Fund/authorize</u> the continuation or expansion of the program

Worksheet 1B
What Matters to Stakeholders

Stakeholders		What activities and/or outcomes of this program matter most to them?
1		
2		
3		
4		
5		
6		
7		
8		

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 1: Engage Stakeholders

Stakeholders for appropriate antibiotic use programs may include:

Those involved in program operations:

- Program managers and staff
- Local, state, and regional coalitions interested in reducing inappropriate antibiotic use
- State and local health departments
- Funding agencies, such as national and state governments

Those served or affected by the program:

- Physicians, nurse practitioners, pharmacists, and other healthcare providers
- Healthcare systems and the medical community
- Managed care organizations and healthcare delivery organizations
- Healthcare insurers and insurer organizations
- Schools and educational groups
- Universities and educational institutions
- Parent Teacher Associations (PTAs)
- Childcare providers and organizations of childcare providers
- Community organizations
- Consumer advocacy groups
- Patients and the general public

The intended users of the evaluation results will vary with each specific evaluation, and often the users comprise a subset of the individuals and groups listed in the prior two categories. Hence, the many potential users of a specific evaluation's results might include:

- Program managers and staff
- Local, state, and regional coalitions interested in reducing inappropriate antibiotic use
- State and local health departments
- Funding agencies, such as national and state governments

Why Stakeholders Matter

Evaluations of appropriate antibiotic use efforts, like evaluations of other public health efforts, will benefit greatly from the involvement of diverse groups of partners and stakeholders. When appropriate antibiotic use programs are planned and implemented by coalitions, coalition members should also be engaged in the planning and implementation of the program's evaluation. Target populations such as patients and providers should be involved in program evaluation to ensure that the evaluation focus meets their needs and that the evaluation is ethical. Engaging target populations in planning evaluation activities will also help ensure that the

evaluation is feasible and accurate. For example, patient questionnaires will yield much better information on patients' knowledge and attitudes if groups of patients have reviewed the questions to make sure they are clear and understandable. Similarly, providers will be far more likely to complete a questionnaire if they have helped design a plan for implementation that does not disrupt clinic flow.

Often, groups of stakeholders will define program success differently; therefore, it is important to understand stakeholders' different interests and expectations from the start. Worksheet 1B can help you determine which components of the program and which outcomes matter most to various stakeholders. Epidemiologists and other health department staff may assume that a "successful" program would result in reductions in antibiotic resistance rates or slower increases in these rates as compared to a control group. Managed care organizations and other health delivery organizations may look to reduced costs as a measure of success (e.g., decreased prescriptions for antibiotics or decreased number of office visits) in addition to improvements in quality of care. Health educators often look at changes in knowledge, attitudes, and behaviors as indicators of success, especially when these intermediate outcomes are quicker to change or easier to measure than more long-term outcomes.

The Role of Stakeholders in Program Evaluation

As discussed earlier, stakeholders can be involved at various levels of program evaluation. Stakeholders can contribute to the program description, suggest or choose evaluation questions, and disseminate evaluation results. Including stakeholders can inspire a change in focus during program planning or program evaluation. For example, a group of healthcare providers may cite high patient demand for antibiotics when describing the problem of antibiotic resistance and inappropriate antibiotic use. Their definition of the problem might lead to a program based on educating consumers about the risks of overuse of antibiotics. If consumers were involved in efforts to define the problem, the resulting program could look quite different. Consumers may say that their medical providers don't listen to their complaints or explain their diagnosis and treatment and that they feel rushed by short office visits. While consumer education would still be an important component, consumer input to the definition of the problem illustrates the need to examine the structure of office visits as well as provider skills in communicating with patients.

Stakeholders can also play an important role in crafting evaluation tools. Healthcare providers can provide useful insight when drafting and selecting evaluation questions for participants in provider education components of appropriate antibiotic use programs. Finally, stakeholders can play key roles in disseminating evaluation results. For example, professional medical societies and managed care organizations can distribute evaluation findings to providers through newsletters, mailings, and other contact with members.

Step 2: Describe the Program

Developing a comprehensive program description is the next step in the CDC Framework. A comprehensive program description clarifies all the components and intended outcomes of the program, thus helping you focus your evaluation on the most central and important questions. Note that in this step you are describing the *program* and not the evaluation. In this chapter, you will use a tool called “logic modeling” to depict these program components, but a program description can be developed without using this or any tool.

This step can either follow the stakeholder step or precede it. In either case, the combination of stakeholder engagement and program description produces clarity and consensus long before data are available to measure program effectiveness. This clarity on activities, outcomes, and their inter-relationships sets the stage for good program evaluation; in addition, they can be helpful in strategic planning and performance measurement, ensuring that insights from these various processes are integrated.

A comprehensive program description includes the following components:

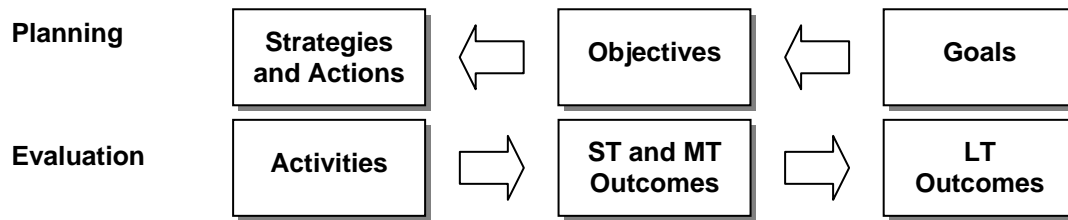
- **Need.** What is the big public health problem you aim to address with your program?
- **Targets.** Which groups or organizations need to change or take action to make progress on the public health problem?
- **Outcomes.** How and in what way do these targets need to change? What action specifically do they need to take?
- **Activities.** What will your program and its staff do to move these target groups to change/take action?
- **Outputs.** What tangible capacities or products will be produced by your program’s activities?
- **Resources/Inputs.** What is needed from the larger environment in order for the activities to be mounted successfully?
- **Relationship of Activities and Outcomes.** Which activities are being implemented to produce progress on which outcomes?

In addition to specifying these components, a complete program description includes discussion of:

- **Stage of Development.** Is the program just getting started, is it in the implementation stage, or has it been underway for a significant period of time?
- **Context.** What factors and trends in the larger environment may influence program success or failure?

Matching Terms from Planning and Evaluation

Planning and evaluation are companion processes. Unfortunately, they tend to use different terms to express similar concepts. This may get confusing and lead to less integration of insights from planning and evaluation than is desirable. As noted in the figure below, plans tend to work from abstract/conceptual goals, then specify the more tangible objectives needed to reach them, and then the strategies needed to reach the objectives. These strategies may be specified as actions, tactics, or a host of other terms. The cross-walk from these planning components to the program description step in an evaluation is relatively straightforward. The strategies will provide insights on the program's activities, the objectives will likely indicate some or all of the target audiences and short-term or intermediate outcomes, and the goal is likely to be close to the long-term outcome desired by the program.



You need not start from scratch in defining the components of your program description. For example, a good source for generating a list of outcomes is the goals and objectives that may already exist for the program in its mission, vision, or strategic plan (see text box). The specific objectives outlined in documents like *Healthy People 2010* are another starting point for defining some components of the program description for public health efforts (see <http://www.health.gov/healthypeople>).

Illustrating Program Descriptions

Let's use some of our cases to illustrate the components of a program description.

Need for the Program

The need is the public health or other problem addressed by the program. You might define the need in terms of its consequences for the state or community, the size of the problem overall, the size of the problem in various segments of the population, and/or significant changes or trends in incidence or prevalence.

For example, the problem addressed by the affordable housing program is compromised life outcomes for low-income families due to lack of stability and quality of housing environments. The problem need for the Childhood Lead Poisoning Prevention (CLPP) program is halting the developmental slide that occurs in children with elevated blood-lead levels (EBLL).

Target Groups

Target groups are the various audiences that the program needs to move into action in order to make progress on the public health problem. For the affordable housing program, action of some kind needs to be taken by eligible families, volunteers, and funders/sponsors. For the CLPP

program, reducing EBLL requires some action by families, healthcare providers, and housing officials, among others.

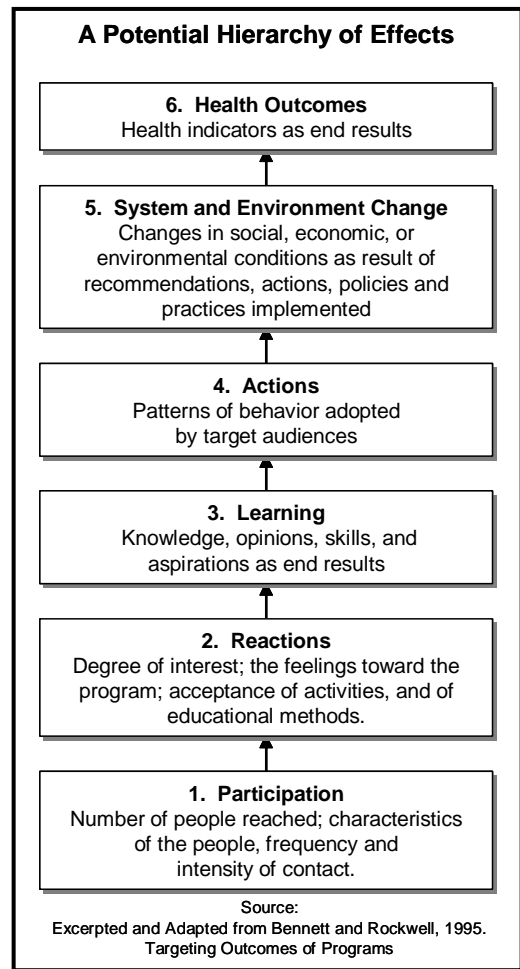
Outcomes

Outcomes²⁶ are the changes in someone or something (other than the program and its staff) that you hope will result from your program’s activities. For programs dealing with large and complex public health problems, the ultimate outcome is often an ambitious and long-term one, such as eliminating the problem or condition altogether or improving the quality of life of people already affected. Hence, a strong program description usually provides details not only on the intended long-term outcomes but on the short-term and intermediate outcomes that precede it and the sequence in which they are likely to occur.

The text box “A Potential Hierarchy of Effects” outlines a potential sequence for a program’s outcomes (effects). Starting at the base of the hierarchy: Program activities aim to obtain *participation* among targeted communities. Participants’ *reactions* to program activities affect their *learning*—their knowledge, opinions, skills, and aspirations. Through this learning process, people and organizations take *actions* that result in a *change in social, behavioral, and/or environmental* condition that directs the long-term *health outcomes* of the community.

In thinking about this hierarchy or any sequence of outcomes, keep in mind that the higher order outcomes are usually the “real” reasons the program was created, even though the costs and difficulty of collecting evidence increase as you move up the hierarchy. Evaluations are strengthened by showing evidence at several levels of hierarchy; information from the lower levels helps to explain results at the upper levels, which are longer term.

The sequence of outcomes for the affordable housing program is relatively simple: Families, sponsors, and volunteers must be engaged and work together for several weeks to complete the house, then the sponsor must sell the house to the family, and then the family must maintain the house payments. For the CLPP program, there are streams of outcomes for each of the target groups: Providers must be willing to test, treat, and refer EBLL children. Housing officials must be willing to clean up houses that have lead paint, and families must be willing to get children and houses screened,



²⁶ Program evaluation and planning are replete with terms that are used inconsistently. In this document, the term “outcomes” is used to refer to the intended changes that will result from the program. However, others may use different terms to refer to the early and late outcomes: results, impacts, and outcomes is a typical sequence.

adopt modest changes in housekeeping behavior, and adhere to any treatment schedule to reduce EBLL in children. Together, these ensure higher order outcomes related to reducing the EBLL and arresting the developmental slide.

Activities

These are the actual actions mounted by the program and its staff to achieve the desired outcomes in the target groups. Obviously, activities will vary with the program. Some typical program activities may include, among others, outreach, training, funding, service delivery, collaborations and partnerships, and health communication. For example, the affordable housing program must recruit, engage, and train the families, sponsors, and volunteers, and also oversee construction and handle the mechanics of home sale. The CLPP program does outreach and screening of children, and, for those children with EBLL, does case management, referral to medical care, assessment of the home, and referral of lead-contaminated homes for cleanup.

Outputs

Outputs are the direct products of activities, usually some sort of tangible deliverable produced as a result of the activities. Outputs can be viewed as activities redefined in tangible or countable terms. For example, the affordable housing program's activities of engaging volunteers, recruiting sponsors, and selecting families have the corresponding outputs: number of volunteers engaged, number of sponsors recruited and committed, and number and types of families selected. The CLPP activities of screening, assessing houses, and referring children and houses would each have a corresponding output: the number of children screened and referred, and the number of houses assessed and referred.²⁷

Resources/Inputs

These are the people, money, and information needed—usually from others outside the program—to mount program activities effectively. It is important to include inputs in the program description because accountability for resources to funders and stakeholders is often a focus of evaluation. Just as important, the list of inputs is a reminder of the type and level of resources on which the program is dependent. If, in fact, intended outcomes are not being achieved, the resources/inputs list reminds you to look there for one reason that program activities could not be implemented as intended.

In the affordable housing program, for example, a supply of supervisory staff, community relationships, land, and warehouse are all necessary inputs to activities. For the CLPP program, funds, legal authority to screen children and houses, trained staff, and relationships with organizations responsible for the activities that the program cannot undertake—in this case,

²⁷ In trying to distinguish “outputs” from “outcomes,” remember that an outcome is a change in someone or something other than the program and its staff. But also remember that these definitions are guidelines and are not set in stone. Often, there are “gray areas” where something might be classified as an output by some programs and an outcome by others. For example, the number of trainees attending my program is an outcome in the sense that someone other than my program staff—the trainee—took an intentional action (attending the training), but many might classify this an output—number of trainees attending—since there really has not been a change in the trainee.

medical treatment and clean-up of homes—are necessary inputs to mount a successful CLPP program.

Stages of Development

Programs can be roughly classed into three stages of development: planning, implementation, and maintenance/outcomes achievement. As will be seen, the stage of development plays a central role in setting a realistic evaluation focus in the next step. A program in the planning stage will focus its evaluation on a very different part of the program than will a program that has been in existence for several years.

For example, both the affordable housing and CLPP programs have been in existence for several years and can be classed in the maintenance/outcomes achievement stage. Therefore, an evaluation of these programs would probably focus on the degree to which outcomes have been achieved and the factors facilitating or hindering the achievement of outcomes.

Context

The context is the larger environment in which the program is immersed. Because external factors can present both opportunities and roadblocks, you should be aware of and understand them. Program context includes politics, funding, interagency support, competing organizations, competing interests, social and economic conditions, and history (of the program, agency, and past collaborations).

For the affordable housing program, some contextual issues are the widespread beliefs in the power of home ownership and in community-wide person-to-person contact as the best ways to transform lives. At the same time, gentrification in low-income neighborhood drives real estate prices up, which can make some areas unaffordable for the program. And some communities, while approving of affordable housing in principle, may resist construction of these homes in their neighborhood. For the CLPP program, some contextual issues include increasing demands on the time and attention of primary healthcare providers, the concentration of EBLL children in low-income and minority neighborhoods, and increasing demands on housing authorities to ameliorate environmental risks.

A realistic and responsive evaluation will be sensitive to a broad range of potential influences on the program. An understanding of the context also lets users interpret findings accurately and assess the findings' generalizability. For example, the affordable housing program might be successful in a small town, but may not work in an inner-city neighborhood without some adaptation.

Relating Activities and Outcomes: Developing and Using Logic Models

Once the components of the program description have been identified, a visual depiction is often a helpful way to summarize the relationship among any or all of the components. This clarity can help with both strategic planning and program evaluation. While there are other ways to depict these relationships, logic models are a common tool employed by evaluators and the tool described most completely in the CDC Framework.

Logic models are *graphic depictions of the relationship between a program’s activities and its intended outcomes*. Two words in this definition bear emphasizing:

- **Relationship:** Logic models convey not only the activities that comprise the program and the inter-relationship of those activities, but the link between those components and outcomes.
- **Intended:** Logic models depict “intended” outcomes of a program’s activities, rather than reality at any point in time. As the starting point for evaluation and planning, the model serves as an “outcomes roadmap” that shows the underlying logic behind the program, i.e., why it should work. That is, of all activities that could have been undertaken to address this problem, these activities are chosen because, if implemented as intended, they should lead to the outcomes depicted. Over time, evaluation, research, and day-to-day experience will deepen the understanding of what does and does not work, and the model will change accordingly.

- | |
|--|
| <p>Other Names for a Logic Model</p> <ul style="list-style-type: none">• Theory of change• Model of change• Theoretical underpinning• Causal chain• Weight-of-evidence model• Roadmap• Conceptual map• Blueprint• Rationale• Program theory• Program hypothesis |
|--|

The logic model requires no new thinking about the program; rather, it converts the raw material generated in the program description into a picture of the program. The remainder of this chapter provides the steps in constructing and elaborating simple logic models. The next chapter, *Focus the Evaluation Design*, shows how to use the model to identify and address issues of evaluation focus and design.

Constructing Simple Logic Models

A useful logic model can be constructed in a few simple steps, as shown here using the CLPP program for illustration.

Develop a list of activities and intended outcomes. While logic models can include all of the components in the text box, we will emphasize using logic models to gain clarity on the relationship between the program’s activities and its outcomes. There are many ways to develop a list of activities and outcomes that you will incorporate into your model, and indeed you may already have a comprehensive list from the program description. But, to stimulate the creation of a comprehensive list, **any** of the following methods will work.

- Review any information available on the program—whether from mission/vision statements, strategic plans, or key informants—

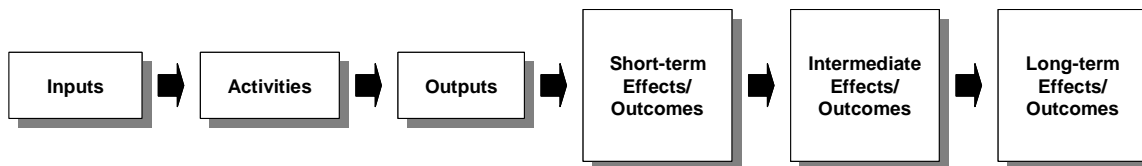
- | |
|--|
| <p>Logic Model Components</p> <p>Logic models may depict all or only some of the following components of your program description, depending on their intended use:</p> <ul style="list-style-type: none">• <u>Inputs:</u> Resources that go into the program and on which it is dependent to mount its activities.• <u>Activities:</u> Actual events or actions done by the program and its staff.• <u>Outputs:</u> Direct products of program activities, often measured in countable terms (e.g., the number of sessions held).• <u>Outcomes:</u> The changes that result from the program’s activities and outputs, often in a sequence expressed as short-term, intermediate, and long-term outcomes. |
|--|

and extract items that meet the definition of activity (something the program and its staff does) and of outcome (some change in someone or something, other than the program and its staff, that you hope will result from the activities), ***or***

- Work backward from outcomes. This is called “reverse logic” logic modeling and may prove helpful when a program is given responsibility for a new or large problem or is just getting started. There may be clarity about the “big change” (most distal outcome) the program is to produce, but little else. Working backward from the distal outcome by asking “how to” will help identify the factors, variables, and actors that will be involved in producing change, ***or***
- Work forward from activities. This is called “forward logic” logic modeling and is helpful when there is clarity about activities but not about why they are part of the program. Moving forward from activities to intended outcomes by asking, “So then what happens?” is often helpful in elaborating downstream outcomes of the activities.

Logic models may depict all or only some of the elements of program description (see text box), depending on the use to which the model is being put. For example, Exhibit 2.1 is a simple, generic logic model. If relevant to the intended use, the model could include references to the remaining components of program description, such as “context” or “stage of development.” Likewise, some of the examples presented below focus mainly on the connection of a program’s activities to its sequence of outcomes. Adding “inputs” and explicit “outputs” to these examples would be a simple matter if needed.

Exhibit 2.1
Basic Program Logic Model



Note that Worksheet 2A at the end of this chapter provides a simple format for doing this categorization of activities and outcomes, no matter what method is used. Here, for the CLPP, we completed the worksheet using the first method.

CLPP Program: Listing Activities and Outcomes	
Activities <ul style="list-style-type: none"> • Outreach • Screening • Case management • Referral to medical treatment • Identification of EBLL children • Environmental assessment • Environmental referral • Family training 	Outcomes <ul style="list-style-type: none"> • Lead source identified • Families adopt in-home techniques • EBLL children get medical treatment • Lead source gets eliminated • EBLL reduced • Developmental “slide” stopped • Quality of Life (Q of L) improved

Subdivide the lists to show the logical sequencing among activities and among outcomes.

Logic models provide clarity on the order in which activities and outcomes are expected to occur. To help provide that clarity, it is useful to take the single column of activities (or outcomes) developed in the last step, and then distribute them across two or more columns to show the logical sequencing. The logical sequencing may be the same as the time sequence, but not always. Rather, the logical sequence says, “Before this activity (or outcome) can occur, this other one has to be in place.”

For example, if the list of activities includes a needs assessment, distribution of a survey, and development of a survey, most would conclude that the needs assessment of content should occur first, and that the distribution of a survey must be preceded by development of the survey. Likewise, among the outcomes, most would generally concede that change in knowledge and attitudes would precede change in behavior.

Worksheet 2B provides a simple format for expanding the initial two-column table. For the CLPP, we expanded the initial two-column table to four columns. Note that no activities or outcomes have been added. But the original lists have been spread over several columns to reflect the logical sequencing. For the activities, we suggest that outreach, screening, and identification of EBLL children need to occur in order to case manage, assess the houses, and refer the children and their houses to follow-up. On the outcomes side, we suggest that outcomes such as receipt of medical treatment, clean-up of the house, and adoption of housekeeping changes must precede reduction in EBLL and elimination of the resultant slide in development and quality of life.

CLPP Program: Sequencing Activities and Outcomes			
Early Activities	Later Activities	Early Outcomes	Later Outcomes
<ul style="list-style-type: none"> • Outreach • Screening • Identification of EBLL children 	<ul style="list-style-type: none"> • Case management • Referral to medical treatment • Environmental assessment • Environmental referral • Family training 	<ul style="list-style-type: none"> • Lead source identified • Lead source gets eliminated • Families adopt in-home techniques • EBLL children get medical treatment 	<ul style="list-style-type: none"> • EBLL reduced • Developmental “slide” stopped • Q of L improved

Add any inputs and outputs. At this point, you may decide that the four-column logic model adds all the clarity that is needed. If not, the next step is often to add columns for inputs and for outputs. The inputs are inserted to the left of the activities while the outputs—as products of the activities—are inserted to the right of the activities but before the outcomes.

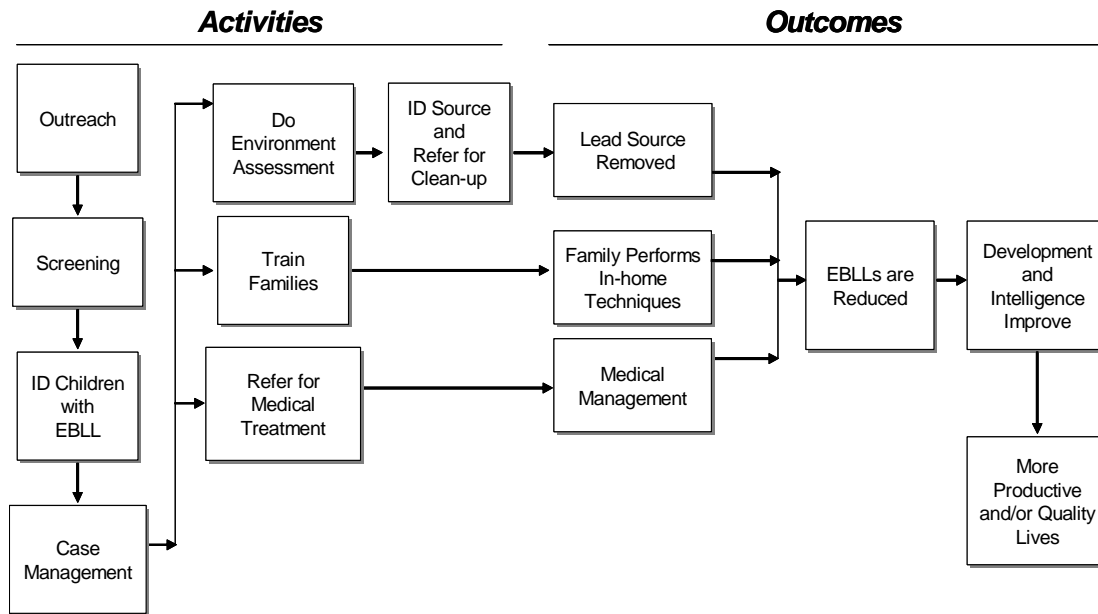
For the CLPP, we can easily define and insert both inputs and outputs of our efforts. Note that the outputs are the products of our activities, but do not confuse them with outcomes. No one has changed yet; while we have identified a pool of leaded houses and referred a pool of EBLL children, the houses have not been cleaned up, nor have the children been treated yet.

CLPP Program: Logic Model with Inputs and Outputs					
Inputs	Early Activities	Later Activities	Outputs	Early Outcomes	Later Outcomes
Funds	Outreach	Case management	Pool (#) of eligible children	Lead source identified	EBLL reduced
Trained staff for screening and clean-up	Screening	Referral to medical treatment	Pool (#) of screened children	Lead source gets eliminated	Developmental “slide” stopped
Relationships with organizations	Identification of EBLL children	Environmental assessment	Referrals (#) to medical treatment	Families adopt in-home techniques	Q of L improved
Legal authority		Environmental referral	Pool (#) of “leaded” homes	EBLL children get medical treatment	
		Family training	Referrals (#) for clean-up		

Draw arrows to depict intended causal relationships. The multi-column table of inputs, activities, outputs, and outcomes that has been developed so far may contain enough detail, depending on the purposes for which the model will be used. In fact, for conveying in a global way the components of a program, it almost certainly will suffice. However, when the model is used to set the stage for planning and evaluation discussions, the logic model will benefit from adding arrows that show the causal relationships among activities and outcomes. These arrows may depict a variety of relationships: from one activity to another, when the first activity exists mainly to feed later activities; from an activity to an outcome, where the activity is intended to produce a change in someone or something other than the program; from an early outcome to a later one, when the early outcome is necessary to achieve the more distal outcome.

Examine the CLPP Logic Model (Exhibit 2.2) with causal arrows included. Note that no activities/outputs or outcomes have been added. Instead, arrows were added to show the relationships among activities and outcomes. Note also that streams of activities exist concurrently to produce cleaned-up houses, medically “cured” children, and trained and active households/families. It is the combination of these three streams that produces reductions in EBLL, which is the platform for stopping the developmental slide and improving the quality of life.

Exhibit 2.2 Lead Poisoning: “Causal” Roadmap



Clean up the logic model. Early versions are likely to be sloppy, and a nice, clean one that is intelligible to others often takes several tries.

Elaborate the Simple Model

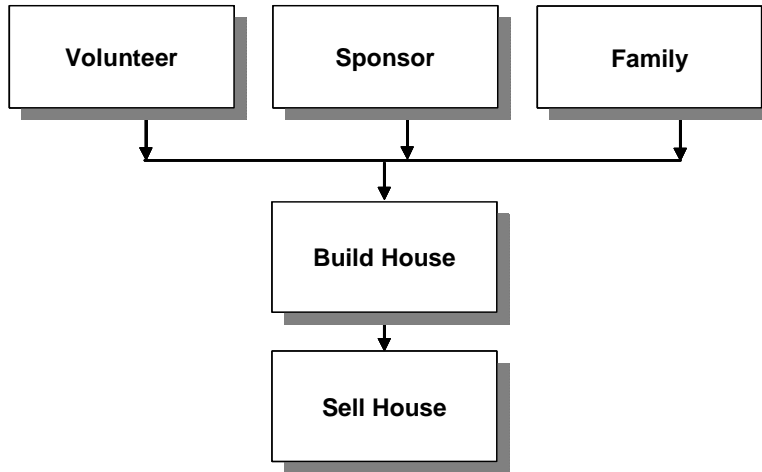
Logic models are a picture depicting your “program theory”—why *should* your program work? The simple logic models developed in these few steps may work fine for that purpose, but often programs benefit from elaborating their simple logic models in some of the following ways:

- **Elaborating distal outcomes:** Sometimes the simple model will end with the short-term outcomes or even outputs. While this may reflect a program’s mission, usually the program has been created to contribute to some larger purpose, and depicting this in the model leads to more productive strategic planning discussions later. This elaboration is accomplished by asking “so then what happens?” of the last outcome depicted in the simple model, and then continuing to ask that of all subsequent outcomes until more distal ones are included.

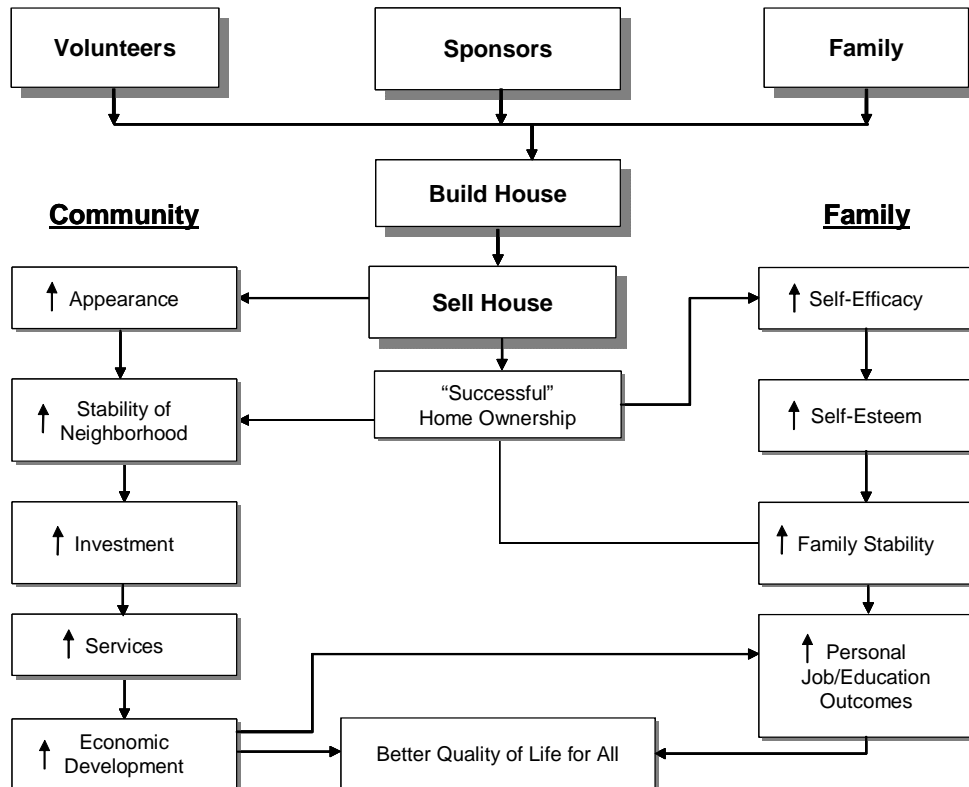
For example, in Exhibit 2.3, the very simple logic model that might result from a review of the narrative about the home ownership program is elaborated by asking, “So then what happens?” Note that the original five-box model remains as the core of the elaborated model, but the intended outcomes now include a stream of more distal outcomes for both the new home-owning families and also for the communities in which houses are built. As will be discussed later, the elaborated model can motivate the organization to think more ambitiously about intended outcomes and whether the right activities are in place to produce them.

**Exhibit 2.3
Elaborating Your Logic Models “Downstream”**

Affordable Housing Program - Logic Model Based on Mission



Affordable Housing Program - Elaborated Logic Model

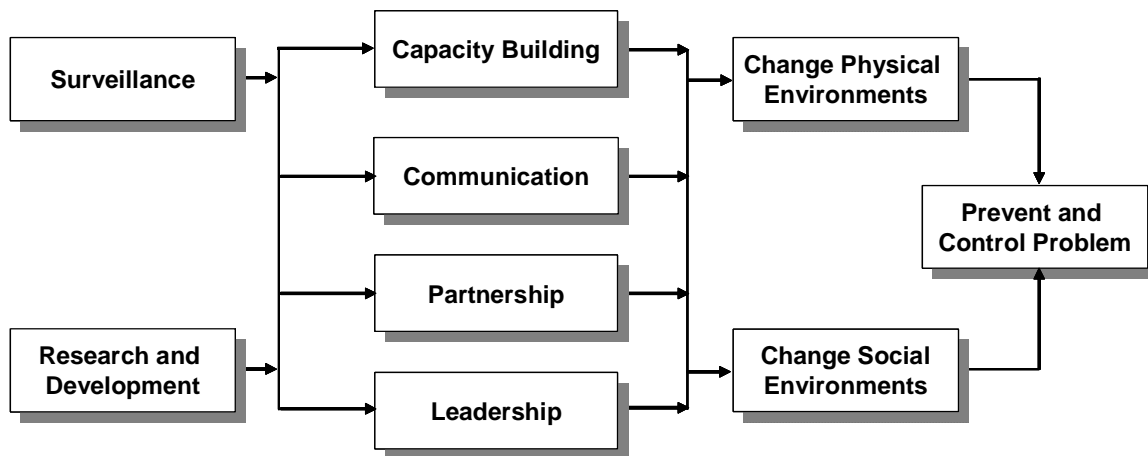


- Elaborating intermediate outcomes:** Sometimes the initial model presents the program’s activities and its most distal outcome in detail, but with scant information on *how* the activities are to produce the outcomes. In this case, the goal of elaboration is to better depict the program logic that links activities to the distal outcomes. Providing such a step-by-step roadmap to a distal destination helps with some or all of the following: identify gaps in program logic that might not otherwise be apparent; persuade skeptics that progress is being made in the right direction, even if the destination has not yet been reached; aid program managers in identifying what needs to be emphasized right now and/or what can be done to accelerate progress.

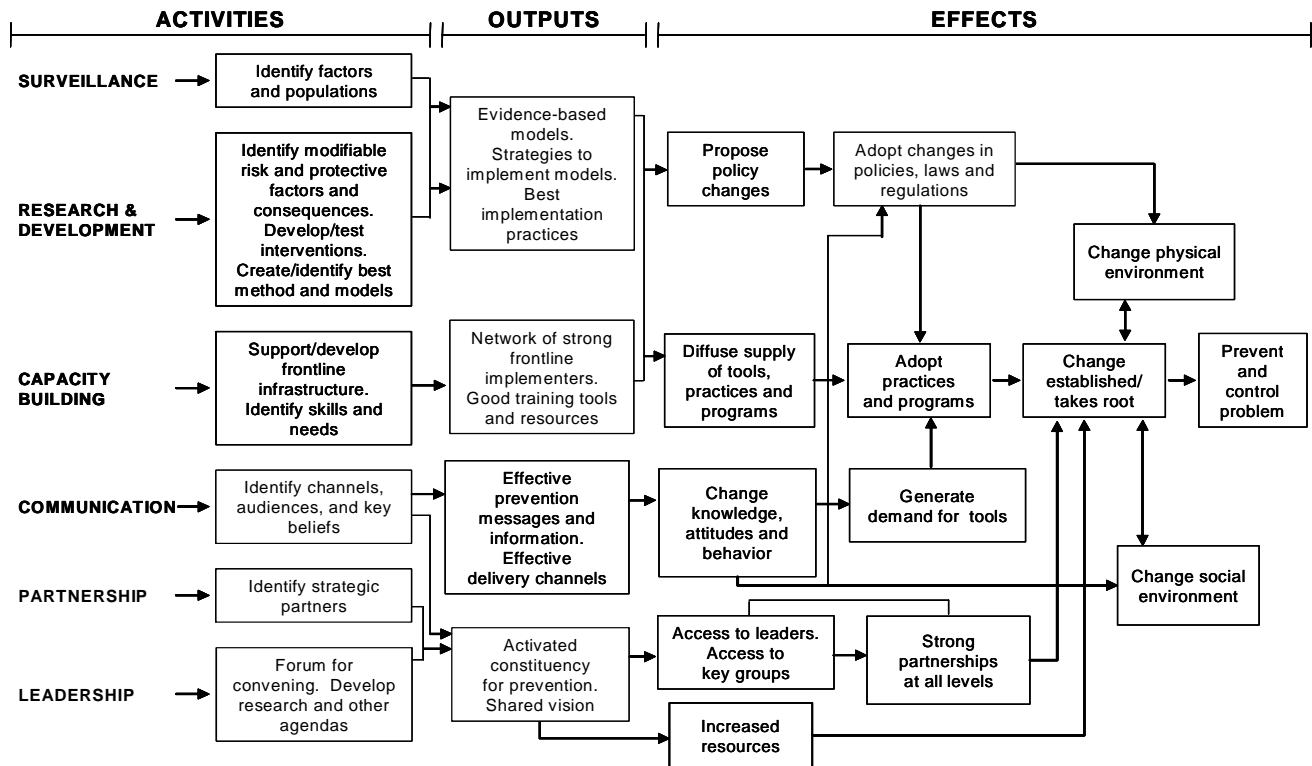
For example, the mission of many CDC programs can be displayed as a simple logic model that shows key clusters of program activities and the key intended changes in a health outcome(s) (Exhibit 2.4). The process of elaboration leads to the more detailed depiction of how the same activities *produce* the major distal outcome, i.e., the milestones along the way.

Exhibit 2.4
Elaborating Intermediate Outcomes in Your Logic Models

Prevention Program - Simple Logic Model



Prevention Program - Elaborated Logic Model



Setting the Appropriate Level of Detail

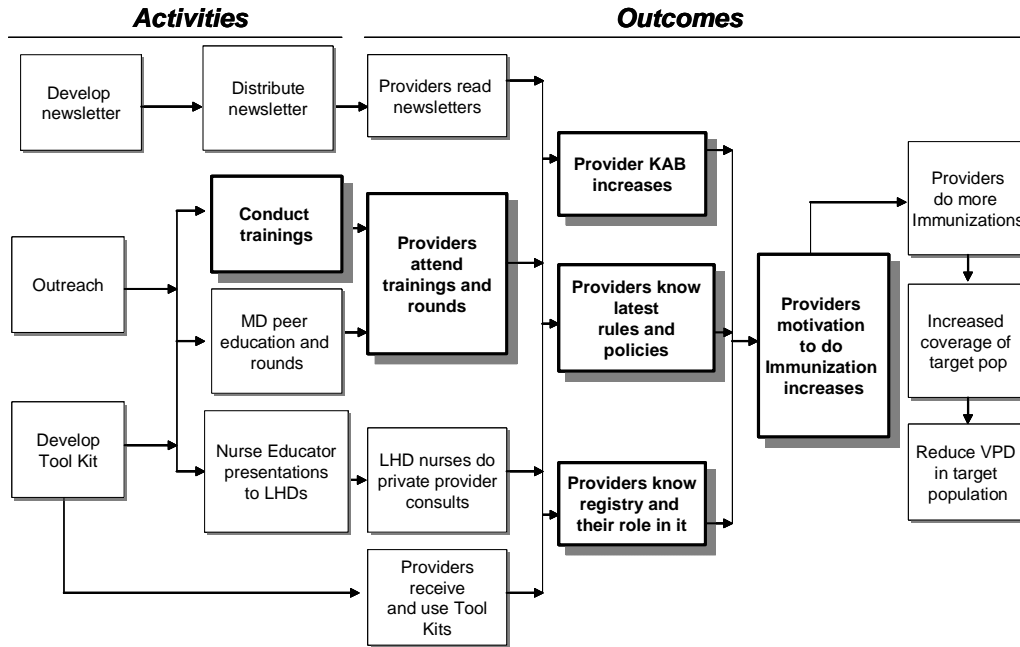
Logic models can be broad or specific. The level of detail depends on the use to which the model is being put and the main audience for the model. A global model works best for stakeholders such as funders and authorizers, but program staff may need a more detailed model that reflects day-to-day activities and causal relationships.

When programs need both global and specific logic models, it is helpful to develop a global model first. The detailed models can be seen as more specific “magnification” of parts of the program. As in geographic mapping programs such as Mapquest, the user can “zoom in” or “zoom out” on an underlying map. The family of related models ensures that all players are operating from a common frame of reference. Even when some staff members are dealing with a discrete part of the program, they are cognizant of where their part fits into the larger picture.

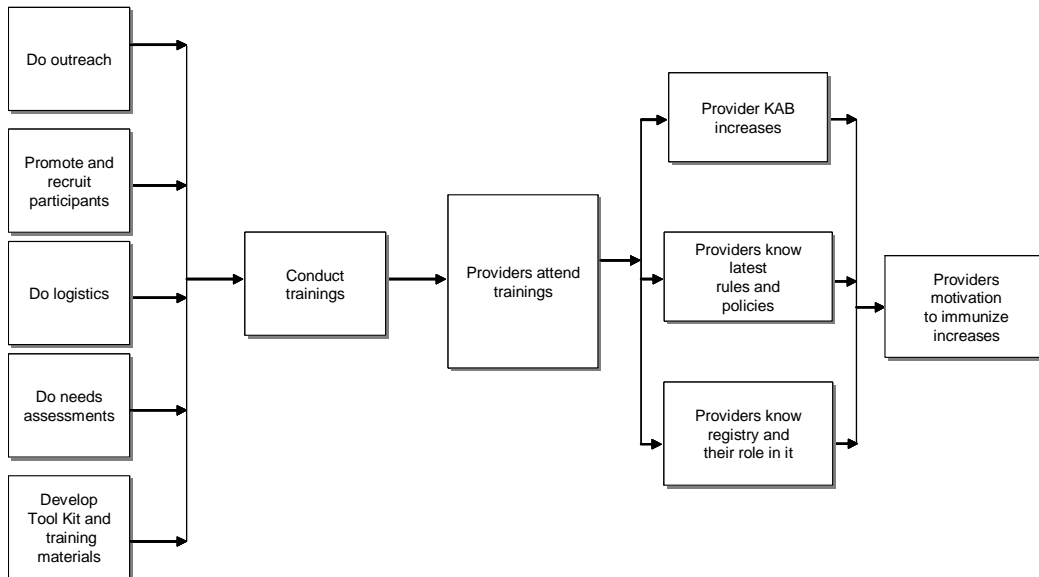
The provider immunization program is a good example of “zooming in” on portions of a more global model. The first logic model (Exhibit 2.5) is a global one depicting all the activities and outcomes, but highlighting the sequence from training activities to intended outcomes of training. The second logic model magnifies this stream only, indicating some more detail related to implementation of training activities.

Exhibit 2.5 Focusing in on Portions of a Program

Provider Education - “Causal” Roadmap - Emphasis on Training Impacts



Provider Education - “Zoom-In” Roadmap - Training



Applying Standards

As in the previous step, you can assure that the evaluation is a quality one by testing your approach against some or all of the four evaluation standards. The two standards that apply most directly to Step 2: Describe the Program are accuracy and propriety. The questions presented in the following table can help you produce the best program description.

Standards for Step 2 Describe the Program

Standard	Questions
Utility	<ul style="list-style-type: none">• Thinking about how the model will be used, is the level of detail appropriate or is there too much or too little detail?• Is the program description intelligible to those who need to use it to make evaluation planning decisions?
Feasibility	<ul style="list-style-type: none">• Does the program description include at least some activities and outcomes that are in control of the program?
Propriety	<ul style="list-style-type: none">• Is the evaluation complete and fair in assessing all aspects of the program, including its strengths and weaknesses?• Does the program description include enough detail to examine both strengths and weaknesses, and unintended as well as intended outcomes?
Accuracy	<ul style="list-style-type: none">• Is the program description comprehensive?• Have you documented the context of the program so that likely influences on the program can be identified?

Checklist for Describing the Program

- Compile a comprehensive program description including need, targets, outcomes, activities, and resources.
- Identify the stage of development and context of the program.
- Convert inputs, activities, outputs, and outcomes into a simple global logic model.
- Elaborate the model as needed.
- Develop more detailed models from the global model as needed.

Worksheet 2A
Raw Material for Your Logic Model

Activities	Outcomes
What will the program and its staff actually do?	What changes do we hope will result in someone or something other than the program and its staff?

Worksheet 2B
Sequencing Activities and Outcomes

Activities		Outcomes	
Early	Later	Early	Later

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 2: Describe the Program

The Need for Appropriate Antibiotic Use Programs

As described earlier, the description of the need for your program should explain the health problem addressed by your program. For appropriate antibiotic use programs, you should answer the following questions:

- Why is antibiotic resistance a health problem? What are the consequences of antibiotic resistance for the state or community?
- What is the size of the problem overall and in various segments of the population?
- What factors contribute to the problem of antibiotic resistance and inappropriate prescribing?
- Who are the target groups for your program?
- What changes or trends are occurring in antibiotic resistance and antibiotic prescribing?

Use local surveillance data when available to show rates of resistant bacteria in your state or community. CDC's Active Bacterial Core surveillance (ABCs) collects and reports data from 10 states (CA, CO, CT, GA, MD, MN, NM, NY, OR, and TN), representing a study population of approximately 38.3 million.²⁸ This information is available at <http://www.cdc.gov/ncidod/dbmd/abcs/>. Check with your state and local health department for the availability of surveillance data in your area. You may also discuss the possibility of conducting surveillance of additional pathogens of interest with your local public health officials. Active, population-based surveillance is resource intensive and not a feasible option for many communities. Aggregating existing hospital antibiograms or cumulative susceptibility data from clinical labs and hospitals is a simpler, less expensive, and relatively accurate way to estimate local resistance rates of drug-resistant *Streptococcus pneumoniae*.²⁹ CDC is developing a surveillance manual to aid state and local health departments in their efforts to monitor resistant bacteria. Sections of this manual are available on-line at <http://www.cdc.gov/DRSPSurveillanceToolKit/index.htm>.

Rates of antibiotic prescribing can also be used to demonstrate the need for your program. Local prescribing rates and rates among various segments of the population are useful to document the need for interventions in your community. Prescribing rates can be analyzed and presented by diagnosis or by antibiotics prescribed to describe current or past inappropriate prescribing practices. Health plans are a rich source of data on antibiotic prescribing.

²⁸ Centers for Disease Control and Prevention. Division of Bacterial and Mycotic Diseases. Active Bacterial Core Surveillance. 2004 Protocol for Active Bacterial Core Surveillance (ABCs) for the Emerging Infection Program Sites. Updated February 2004.

²⁹ Van Beneden CA, Lexau C, Baughman W, et al: Aggregated antibiograms and monitoring of drug-resistant *Streptococcus pneumoniae*. *Emerging Infectious Diseases* 2003;9(9):1089-1095.

When possible, use local data on the determinants of inappropriate antibiotic use and inappropriate prescribing. This could include surveys or focus groups of providers assessing their knowledge of prescribing guidelines, prescribing practices, or perceived pressures to prescribe antibiotics. Surveys or focus groups of consumers assessing knowledge and attitudes surrounding antibiotic use or expectations for antibiotics could also help you describe the problem of inappropriate use in your community.

It is also important to look for disparities among specific populations or communities when discussing the need for your program. Higher rates of antibiotic use have been seen among whites as compared with other racial/ethnic groups, most likely due to their increased healthcare utilization.²⁰ However, specific populations or communities may have particular risk factors for inappropriate antibiotic use. For example, while many immigrant populations have limited access to conventional medical care, some immigrant groups have easy access to antibiotics and other medicines imported from other countries. In addition, historically marginalized groups may view appropriate antibiotic use messages as attempts by the dominant culture to further restrict their access to care. In this case, information about knowledge and attitudes surrounding antibiotic use, access to antibiotics, and usual sources of medicines and health care would all contribute to a greater understanding of the need for your program and the development of appropriate interventions. Sources of data for these indicators include, but are not limited to, national and state surveys, regional or community surveys, case studies, expert panels, and stakeholder panels.

Program Activities

Given the number of options for activities to promote appropriate antibiotic use, it is helpful to mention some existing frameworks and guidelines that can be used to direct the selection of program activities. These include the idea of social ecology as a theoretical basis and a list of recommended components and strategies based on the results of successful efforts. Regardless of the activities selected, a good program description will explain the reason for choosing these activities and will list their intended outcomes.

CDC encourages state and local programs to adopt a social ecological framework when designing interventions to promote appropriate antibiotic use. The social ecological approach to health promotion recognizes the contributions of both individual influences (e.g., knowledge, attitudes, and skills) and social environmental influences (e.g., social norms and organizational policies and practices) on health behavior.^{30,31,32} This approach integrates person-focused efforts to modify people's health behavior with environment-focused interventions to increase support for behavior change in their physical and social surroundings.³³

³⁰ McLeroy KR, Bibeau D, Steckler A, et al: An ecological perspective on health promotion programs. *Health Education Quarterly* 1988;15(4):351-377.

³¹ Green, LW, Richard, L, Potvin, L: Ecological foundations of health promotion. *American Journal of Health Promotion* 1996;10(4):270-281.

³² Corbett, KK: Susceptibility of youth to tobacco: a social ecological framework for prevention. *Respiration Physiology* 2001;128:103-118.

³³ Stokols, D: Translating social ecological theory into guidelines for community health promotion. *American Journal of Health Promotion* 1996;10(4):282-298.

Appropriate antibiotic use programs designed with social ecology in mind would promote changes at both the individual and social environmental levels. Patient and provider education targets individual change in knowledge, attitudes, skills and behavior, or group changes (e.g., social norms). Organizational changes could include revising childcare or workplace policies that require antibiotic treatment before returning after a sickness, or managed care policies for prescribing and pricing of antibiotics.

Based on the results of published trials and extensive experience with other state and local efforts, CDC has some recommendations on components and strategies that appropriate antibiotic use programs might effectively utilize. These components and strategies, some examples of how they are being implemented, and their rationale are shown in Exhibit 2.6.

Exhibit 2.6: Recommended Appropriate Antibiotic Use Program Components and Strategies³⁴

Recommended Program Components / Strategies	Examples	Rationale
Form a coalition of diverse partners.	State and local health departments, health plans, professional provider organizations, medical schools, Parent Teacher Associations (PTAs), school nurses, childcare providers, pharmacies, pharmaceutical manufacturers, and healthcare purchasers.	Partners bring a variety of resources, including staff, time, funding, and knowledge of and access to the target population. Engaging a diverse group of partners can help build community support for the program and expand the program's reach.
Target changes at multiple levels – individuals, groups, and organizations or institutions.	Patient and provider education targets changes in individuals (e.g., knowledge, attitudes, and behavior) and in groups (e.g., social norms). Organizational changes may include revising childcare policies excluding children who have not received antibiotics or workplace policies requiring antibiotic treatment before an employee can return to work.	Interventions focusing on individual change alone may promote a victim-blaming ideology and neglect the importance of social influences on health and disease. ³⁵
Educate providers	Distribution of guidelines; patient education materials; articles in local health journals; grand rounds, lectures, workshops and other CME events; physician-led “peer” education; prescribing rate feedback.	Declines in antibiotic prescribing associated with patient and provider education components were not seen with a limited patient education intervention. ³⁶

³⁴ Weissman J, Besser RE: Promoting appropriate antibiotic use for pediatric patients: a social ecological framework. *Seminars in Pediatric Infectious Disease* 2004;15(1):41-51.

³⁵ McLeroy KR, Bibeau D, Steckler A, et al: An ecological perspective on health promotion programs. *Health Education Quarterly* 1988;15(4):351-377.

³⁶ Gonzales R, Steiner JF, Lum A, et al: Decreasing antibiotic use in ambulatory practice: impact of a multidimensional intervention on the treatment of uncomplicated acute bronchitis in adults. *JAMA* 281:1512-9, 1999.

Educate patients	Posters and pamphlets in waiting rooms; household mailings to patients; “cold kits” containing over-the-counter remedies for symptomatic relief.	Physicians report parental pressure to prescribe antibiotics for their children. ³⁷ In one survey, educating parents was cited by 78% of physician respondents as the single most important program for reducing inappropriate antibiotic use. ³⁸
Educate the public	Educational materials distributed in community settings (i.e., schools, day care, health fairs); TV, radio and newspaper coverage. CDC national media campaign materials and toolkit are available for use by local campaigns.	Multiple exposures to program messages are more likely to produce changes in knowledge, attitudes and behavior. Media efforts can help raise awareness and change community norms. National media efforts can build national agenda and create foundation for local efforts. ³⁹
Evaluate program efforts	Program evaluation of both process and outcome measures.	Program evaluation can be used to: monitor progress toward the program’s goals, demonstrate that a particular program or activity is effective, identify activities that are ineffective and learn how to improve programs, justify the need for further funding and support, and communicate with stakeholders. ⁴⁰

Program Outcomes

Outcomes are the changes in someone or something (other than the program and its staff) that we hope will result from program activities. The goals and objectives for a program will help guide the determination of intended program outcomes. Programs can also look to national measures when selecting target outcomes for local programs. Both Healthy People 2010 and the Health Plan Employer Data and Information Set (HEDIS[®]), the performance measurement tool used by over 90 percent of the nation’s health plans, include measures on appropriate antibiotic use that can be used in evaluating local programs.

Healthy People 2010 includes two objectives that measure the appropriate use of antibiotics. The first objective measures the use of antibiotics for ear infections among children less than five years old, and the second objective measures the use of antibiotics for the common cold. See <http://www.healthypeople.gov/> for more information.

³⁷ Barden LS, Dowell SF, Schwartz B, et al: Current attitudes regarding use of antimicrobial agents: Results from physicians’ and parents’ focus group discussions. *Clinical Pediatrics* 1988;37:665-672.

³⁸ Bauchner H, Pelton, SI, Klein, JO: Parents, physicians, and antibiotic use. *Pediatrics* 1999;103(2):395-401.

³⁹ Finnegan JR, Viswanath K: Communication theory and health behavior change: the media studies framework, in Glanz K, Lewis FM, Rimer BK (eds): *Health Behavior and Health Education: Theory, Research, and Practice* (ed 2). San Francisco, CA, Jossey-Bass Inc, 1999, pp 313-341.

⁴⁰ MacDonald G, Starr G, Schooley M, et al: *Introduction to Program Evaluation for Comprehensive Tobacco Control Programs*. Atlanta, GA, Centers for Disease Control and Prevention, 2001.

CDC and the National Committee on Quality Assurance have written two pediatric and two adult measures for HEDIS[®]. The pediatric measures, which were incorporated into HEDIS[®] in 2004, assess the appropriate treatment of children who present with pharyngitis or upper respiratory infections. The pharyngitis measure calculates the proportion of children who are tested for group A strep before receiving antibiotics for sore throats. The upper respiratory tract measure looks at the proportion of children who do not receive an antibiotic for the common cold. The adult measures, which were incorporated into HEDIS[®] in 2006 and complement the pediatric measures, assess inappropriate antibiotic treatment of adults with acute bronchitis as well as outpatient utilization of antibiotic prescriptions. The acute bronchitis measure calculates the percentage of healthy adults 18–64 years of age with a diagnosis of acute bronchitis who were dispensed an antibiotic prescription on or three days after the episode date. This Effectiveness of Care process measure assesses if antibiotics were inappropriately prescribed for healthy adults with bronchitis. The outpatient utilization of antibiotic prescriptions measure, which is a Use of Services measure, summarizes the following data: total number of antibiotic prescriptions, average number of antibiotic prescriptions per member per year, total days supplied for all antibiotics, average number of days supplied per antibiotic prescription, average number of antibiotics per member per year and reported by drug class (for selected antibiotics of concern; for all other antibiotics), and percentage of antibiotics of concern of total antibiotic prescriptions (during the measurement year, stratified by age and gender, and reported for each product). See <http://www.ncqa.org/communications/publications/hedispub.htm> for more information.

For appropriate antibiotic use, intended program outcomes typically include changes in the knowledge, attitudes, or behaviors of either patients, providers or the general public, depending on the focus of program activities. More specifically, patient education activities can aim to increase patients' knowledge of appropriate antibiotic use messages, change their attitudes to support appropriate antibiotic use, decrease their likelihood of demanding antibiotics from their providers, or increase their adherence to antibiotics when prescribed. For providers, educational activities can aim to increase providers' knowledge about appropriate prescribing, change their attitudes and norms to support appropriate prescribing, or increase appropriate prescribing behavior. In addition, program activities may target changes at the organizational level, such as policies at the workplace, in childcare settings, or in managed care organizations. In these cases, outcomes would include the implementation of policies that support appropriate prescribing.

These outcomes can be divided into short-term, intermediate, and long-term outcomes, depending on the relative length of time needed to achieve change. Changes in knowledge, attitudes, and skills are relatively easy to accomplish and are usually classified as short-term outcomes. Behavior changes, such as patient demand for antibiotics or provider adherence to prescribing guidelines, are more difficult to achieve and would be classified as either intermediate or long-term outcomes.

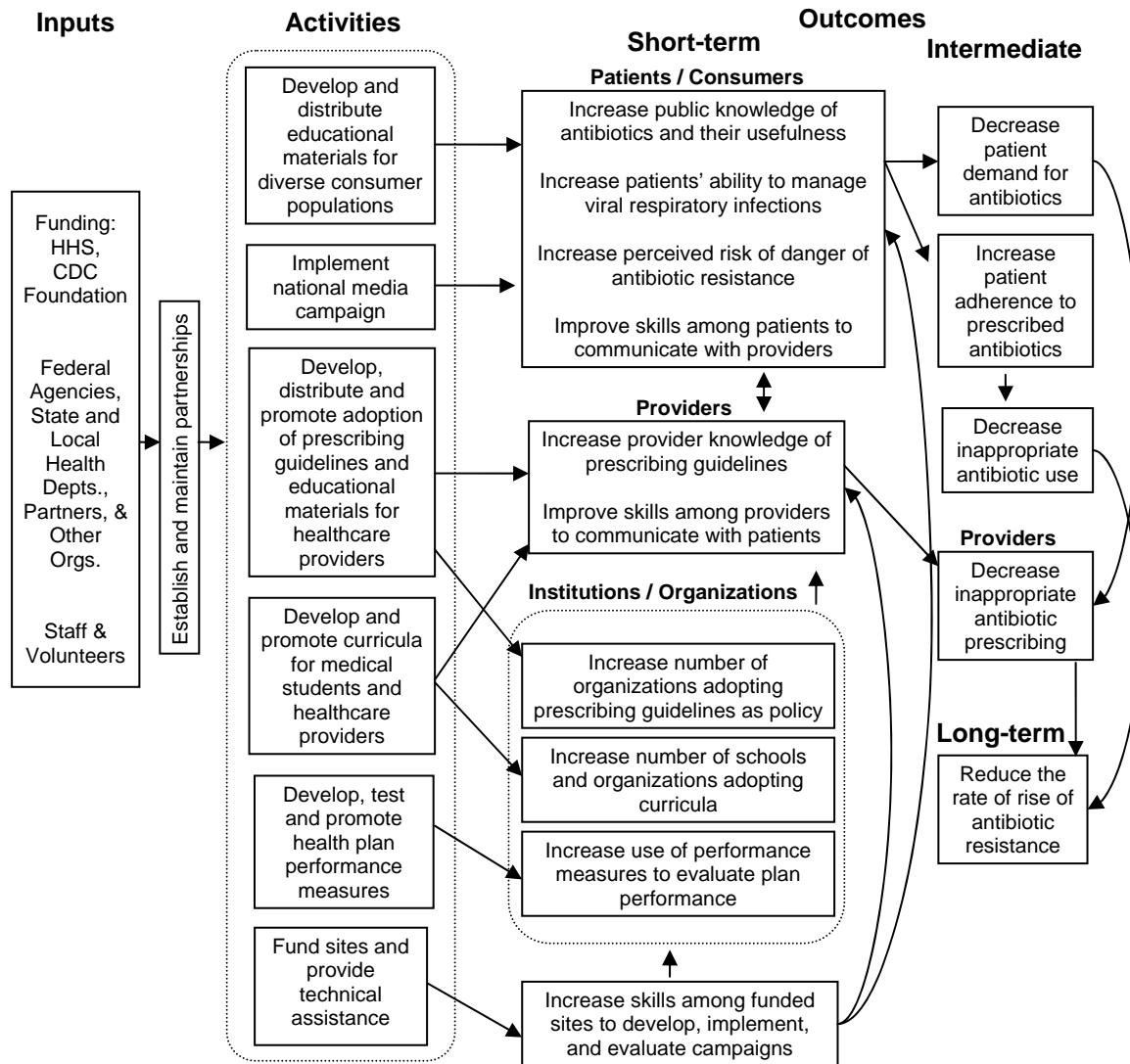
Logic Models

As described earlier, logic models are graphic depictions (i.e., pictures) of the relationship between a program's activities and its intended effects. In addition to presenting a clear and succinct picture of the program components and their intended results, the program's logic model

is used in program evaluation to identify performance indicators and to help in the selection of the activities and outcomes that will be included in any given evaluation.

Two logic models are presented here as examples. The first of the following logic models depicts CDC's Get Smart: Know When Antibiotics Work program activities at the national level (Exhibit 2.7). Providing funds and technical assistance for state and local programs is a key activity of CDC's national program, and the relationship of this activity to other national activities and to their intended outcomes is shown here. State and local campaigns and their intended effects are further depicted in the second logic model (Exhibit 2.8).

Exhibit 2.7: Logic Model, *Get Smart: Know When Antibiotics Work* – National Activities



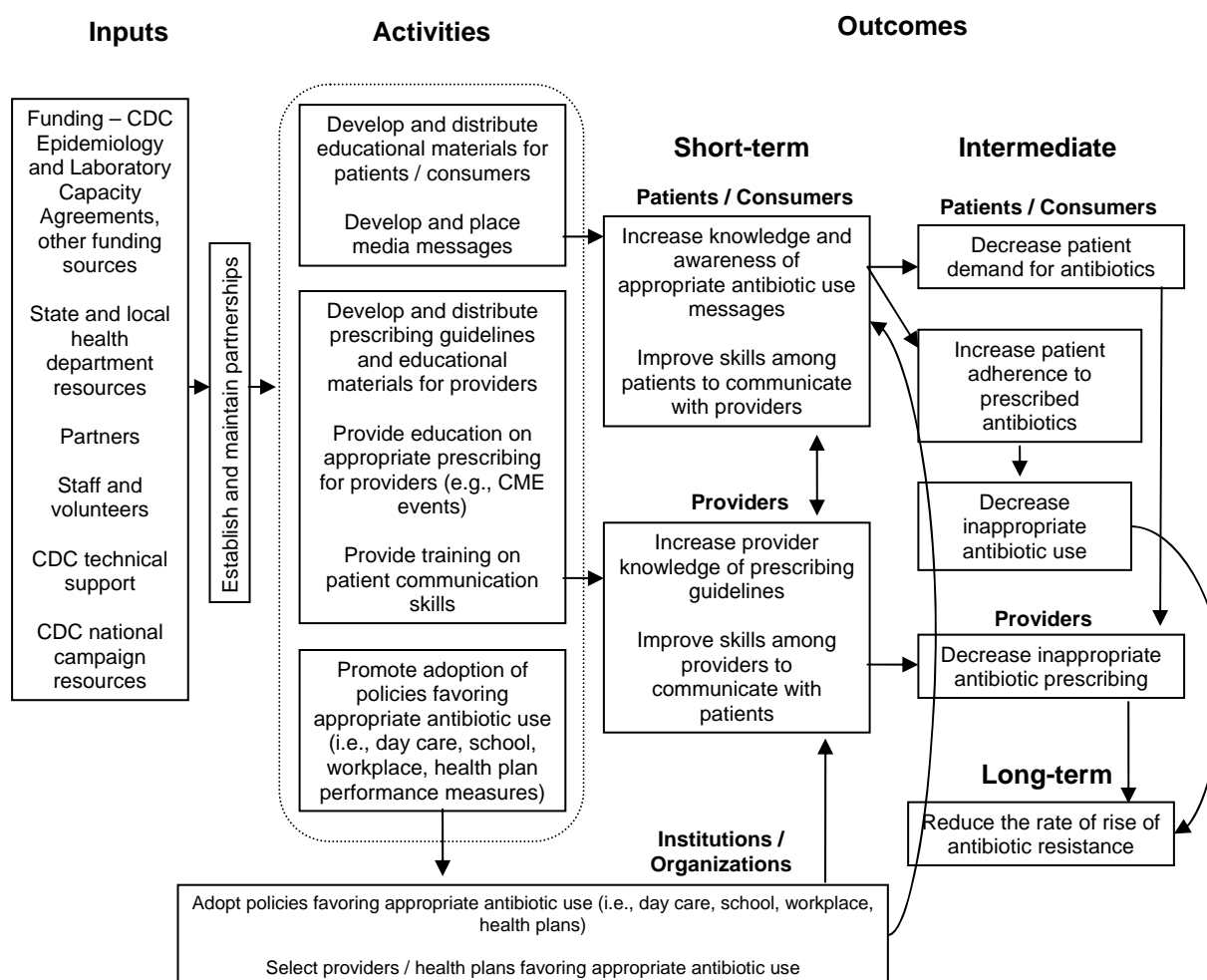
***Get Smart: Know When Antibiotics Work* – National Activities**

In this model, key inputs include staff, partners (including federal agencies, health departments, and other organizations), and funding, which allow the implementation of a variety of activities through the establishment and maintenance of partnerships.

Key activities of the national campaign are listed in this model, and the relationship between these activities and their intended effects is shown. Outcomes are grouped by their respective target audiences – patients/consumers, providers, or institutions/organizations. For example, short-term outcomes for patients or consumers include increases in knowledge, perceived risk, communication skills, and the ability to manage infections. These in turn lead to intermediate outcomes of decreased demand for antibiotics, increased adherence to prescribed antibiotics, and

decreased antibiotic use, which ultimately are expected to contribute to a reduction in the rate of rise of antibiotic resistance. Anticipated relationships between outcomes for different target audiences are also depicted here. For example, decreased inappropriate prescribing by providers is achieved not only through educating providers and subsequent increases in their knowledge and skills as might be expected, but also through educating patients and decreasing their demand for antibiotics, promoting the adoption of prescribing guidelines and curricula, and promoting the use of performance measures.

Exhibit 2.8: Logic Model, State and Local Appropriate Antibiotic Use Campaigns



State and Local Appropriate Antibiotic Use Campaigns

This model provides more detail for the state and local activities referred to in the global model depicting Get Smart’s national activities. Key inputs include staff, partners, funding (including the Epidemiology and Laboratory Capacity cooperative agreements), as well as other health department resources, CDC technical support, and CDC national campaign resources.

This model is meant to generally illustrate the types of activities implemented and outcomes expected by state and local campaigns. Because program staff and coalitions develop and implement activities in response to local needs and resources, CDC expects these efforts to vary across communities. Typical activities for state and local campaigns are listed in this model, and again, outcomes are grouped by their respective target audiences – patients/consumers, providers, or institutions/organizations. Of course, actual outcomes will vary depending on the activities implemented.

While the “right” logic model will vary with each situation, as mentioned earlier, elaborating a simple logic model can prove useful in program planning or evaluation. A simple initial model detailing appropriate antibiotic use activities may only depict the program’s short-term outcomes. This type of model may adequately guide your work and much of your evaluation even though it does not show the ultimate goal of your program. For example, a simple logic model for a media campaign may show posters, brochures, and public service announcements leading to an increase in public knowledge and awareness about appropriate antibiotic use. Critics of this program may question whether changes in knowledge and awareness alone are a worthy goal. However, if you believe that these increases in knowledge and awareness will help to decrease patient demand for antibiotics and ultimately contribute to more appropriate prescribing (especially in conjunction with education for healthcare providers), including these more distal effects and the relationship between activities and effects will help to explain your choice of short-term outcomes as evaluation measures. In this way, if you are able to make a reasonable case for the relationship between activities and intended effects, you can justify using more proximal outcomes as measures of program success when the long-term outcomes are slow to change and/or difficult to measure.

Step 3: Focus the Evaluation Design

After completing Steps 1 and 2, you and your stakeholders should have a clear understanding of the program and reached consensus. Now your evaluation team will need to focus the evaluation. This includes determining the most important evaluation questions and the appropriate design for the evaluation. Focusing the evaluation is based on the assumption that the entire program does not need to be evaluated at any point in time. Rather, the “right” evaluation of the program depends on what question is being asked, who is asking the question, and what will be done with the information.

Since resources for evaluation are always limited, this chapter provides a series of decision criteria to help you determine the best evaluation focus at any point in time. You will note that these criteria are inspired by the evaluation standards: specifically, utility (who will use the results and what information will be most useful to them) and feasibility (how much time and resources are available for the evaluation).

The logic models developed in the prior step set the stage for determining the best evaluation focus. The approach to evaluation focus in the CDC Evaluation Framework differs slightly from traditional evaluation approaches. In the past, some programs tended to assume all evaluations were “summative” ones, conducted when the program had run its course and intended to answer the question, “Did the program work?” Consequently, a key question was, “Is the program ready for evaluation?”

By contrast, the CDC Framework views evaluation as an ongoing activity over the life of a program that asks, “Is the program working?” Hence, a program is always ready for *some* evaluation. Because the logic model displays the program from inputs through activities/outputs through to the sequence of outcomes from short-term to most distal, it can guide a discussion of what you can expect to achieve at this point in the life of your project. Should you focus on distal outcomes, or only on short- or mid-term ones? Or conversely, does a process evaluation make the most sense right now?

Types of Evaluations

Many different questions can be part of a program evaluation, depending on how long the program has been in existence, who is asking the question, and why the evaluation information is needed. In general, evaluation questions for an existing program⁴¹ fall into one of the following groups:

⁴¹ There is another type of evaluation—“formative” evaluation—where the purpose of the evaluation is to gain insight into the nature of the problem so that you can “formulate” a program or intervention to address it. While many steps of the Framework will be helpful for formative evaluation, the emphasis in this manual is on instances wherein the details of the program/intervention are already known even though it may not yet have been implemented.

Implementation/Process

Implementation evaluations (more commonly called “process evaluations”) document whether a program has been implemented as intended—“implementation fidelity”—and why or why not? In conducting process evaluations, you might examine whether the activities are taking place, who is conducting the activities, who is reached through the activities, and whether sufficient inputs have been allocated or mobilized. Process evaluation is important to help you distinguish the causes of poor program performance—was the program a bad idea, or was it a good idea that could not reach the standard for implementation that you set? In all cases, process evaluations measure whether actual program performance was faithful to some initial plan. This might include contrasting actual and planned performance on all or some of the following:

- 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

When evaluation resources are limited, only the most important issues of implementation fidelity can be included. Here are some “usual suspects” that compromise implementation fidelity and should be considered for inclusion in the process evaluation portion of the evaluation focus:

- **Transfers of Accountability:** Where a program’s activities cannot produce the intended outcomes unless some other person or organization takes appropriate action, there is a transfer of accountability.
- **Dosage:** The intended outcomes of program activities (e.g., training, case management, counseling) may presume a threshold level of participation or exposure to the intervention.
- **Access:** Where intended outcomes require not only an increase in consumer demand but also an increase in supply of services to meet it, then the process evaluation might include measures of access.
- **Staff Competency:** The intended outcomes may presume well-designed program activities that are delivered by staff who are not only technically competent but also are matched appropriately with the target audience. Measures of the match of staff and target audience might be included in the process evaluation.

Our childhood lead poisoning logic model illustrates many of these potential process issues. Reducing EBLL presumes the house will be cleaned, medical care referrals will be fulfilled, and specialty medical care will be provided. All of these are transfers of accountability beyond the program to the housing authority, the parent, and the provider, respectively. For provider training to achieve its outcomes, it may presume completion of a three-session curriculum, which

is a dosage issue. Case management results in medical referrals, but it presumes adequate access to specialty medical providers. And because lead poisoning tends to disproportionately affect children in low-income urban neighborhoods, many program activities presume cultural competence of the caregiving staff. Each of these components might be included in a process evaluation of a childhood lead poisoning prevention program.

Effectiveness/Outcome

Outcome evaluations assess progress on the sequence of outcomes that the program is to address. Programs often describe this sequence using terms like short-term, intermediate, and long-term outcomes, or proximal (close to the intervention) or distal (distant from the intervention). Depending on the stage of development of the program and the purpose of the evaluation, outcome evaluations may include any or all of the outcomes in the sequence, including

- Changes in people's attitudes and beliefs
- Changes in risk or protective behaviors
- Changes in the environment, including public and private policies, formal and informal enforcement of regulations, and influence of social norms and other societal forces
- Changes in trends in morbidity and mortality.

While process and outcome evaluations are the most common, there are several other types of evaluation questions that are central to a specific program evaluation. These include the following:

Efficiency: Are your program's activities being produced with minimal use of resources such as budget and staff time? What is the volume of outputs produced by the resources devoted to your program?

Cost-Effectiveness: Does the value or benefit of your program's outcomes exceed the cost of producing them?

Attribution: Can the outcomes that are being produced be shown to be related to your program, as opposed to other things that are going on at the same time?

All of these types of evaluation questions relate to some part, but not all, of the logic model. Exhibits 3.1 and 3.2 show where in the logic model each type of evaluation would focus. Implementation evaluations would focus on the inputs, activities, and outputs boxes and not be concerned with performance on outcomes. Effectiveness evaluations would do the opposite—focusing on some or all *outcome boxes*, but not necessarily on the activities that produced them. Efficiency evaluations care about the *arrows linking inputs to activities/outputs*—how much output is produced for a given level of inputs/resources. Attribution would focus on the *arrows between specific activities/outputs and specific outcomes*—whether progress on the outcome is related to the specific activity/output.

Exhibit 3.1
Evaluation Domains — Boxes

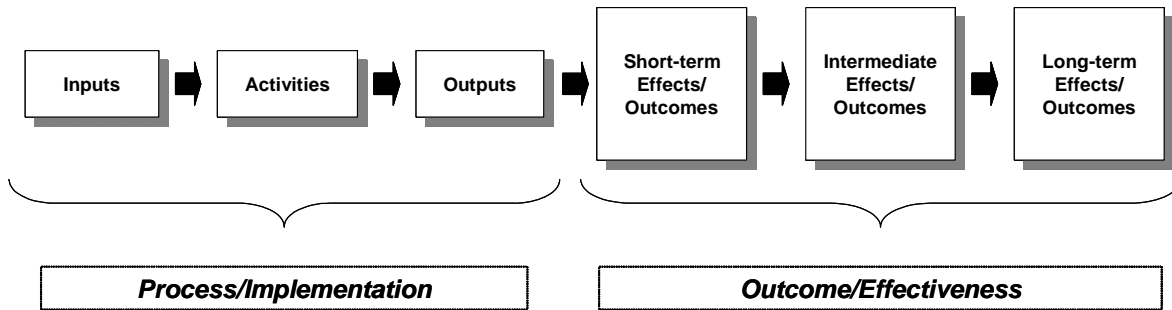
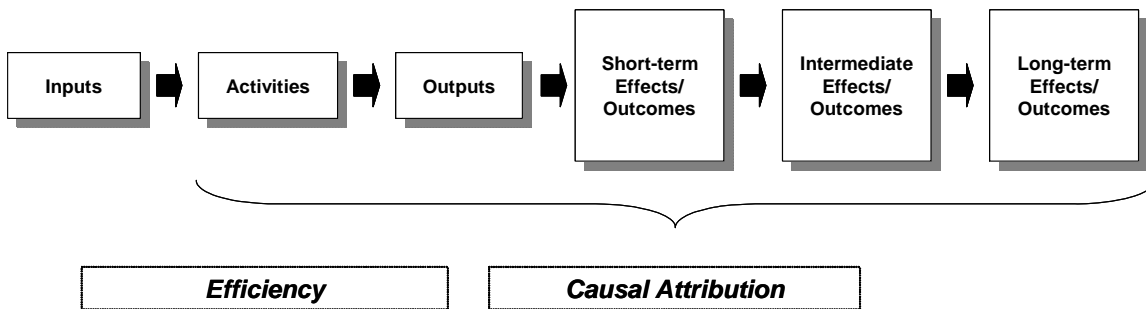


Exhibit 3.2
Evaluation Domains — Arrows



Determining the Evaluation Focus

Determining the correct evaluation focus is a case-by-case decision. As noted, several guidelines inspired by the “utility” and “feasibility” evaluation standards can help determine the best focus.

Utility Considerations

1) What is the purpose of the evaluation?

Purpose refers to the general intent of the evaluation. A clear purpose serves as the basis for the evaluation questions, design, and methods. Some common purposes:

- Gain new knowledge about program activities
- Improve or fine-tune existing program operations (e.g., program processes or strategies)
- Determine the effects of a program by providing evidence concerning the program’s contributions to a long-term goal
- Affect program participants by acting as a catalyst for self-directed change (e.g., teaching)

2) Who will use the evaluation results?

Users are the individuals or organizations that will employ the evaluation findings in some way. The users will likely have been identified during Step 1 during the process of engaging stakeholders. In this step, you need to secure their input into the design of the evaluation and the selection of evaluation questions. Support from the intended users will increase the likelihood that the evaluation results will be used for program improvement.

3) How will they use the evaluation results?

Uses describe what will be done with what is learned from the evaluation, and many insights on use will have been identified in Step 1. Information collected may have varying uses, which should be described in detail when designing the evaluation. Some examples of uses of evaluation information:

- To document the level of success in achieving objectives
- To identify areas of the program that need improvement
- To decide how to allocate resources
- 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
- To solicit more funds or additional partners

4) What do *other* key stakeholders need from the evaluation?

Of course, the most important stakeholders are those who are requesting or who will use the evaluation results. Nevertheless, in Step 1, you may also have identified stakeholders who, while they are not the users of the findings of the current evaluation, have key questions that may need to be addressed in the evaluation to keep them engaged. For example, a particular stakeholder may always be concerned about costs, disparities, or attribution. If so, and if that stakeholder is important long-term to credibility, implementation, or funding, then you may need to add those questions to your evaluation focus.

Feasibility Considerations

The first four questions help identify the most useful focus of the evaluation, but you must also determine whether it is a realistic/feasible one. Three questions provide a “reality check” on our desired focus:

5) What is the stage of development of the program?

During Step 2, you will have identified the program's stage of development. As noted, there are roughly three stages in program development: planning, implementation, and maintenance. These stages suggest different focuses. In the planning stage, a truly formative evaluation—who is your target, how do you reach them, how much will it cost—may be the most appropriate focus. An evaluation that included outcomes would make little sense at this stage. Conversely, an evaluation of a program in maintenance stage would need to include some measurement of progress on outcomes, even if it also included measurement of implementation.

6) How intensive is the program?

Some programs are wide-ranging and multifaceted. Others may use only one approach to address a large problem. Some programs provide extensive exposure (“dose”) of the program, while others involve participants quickly and superficially. Simple or superficial programs, while potentially useful, cannot realistically be expected to make significant contributions to distal outcomes of a larger program, even when they are fully operational.

7) What are relevant resource and logistical considerations?

Resources and logistics may influence the decision about evaluation focus. Some outcomes are quicker, easier, and cheaper to measure, while others may not be measurable at all. In the short run, at least, these facts may tilt the decision about evaluation focus toward some outcomes as opposed to others.

Early identification of any inconsistencies between “utility” and “feasibility” is an important purpose of the evaluation focus step. For evaluation results to be used, the focus must include questions that matter to those who will implement or otherwise use the results. But we must also ensure a “meeting of the minds” on what is a realistic focus for program evaluation at any point in time.

The affordable housing example shows how the desired focus might be constrained by “reality.” The elaborated logic model was important in this case because it clarified that, while program staff were focused on production of new houses, important stakeholders like community-based organizations and faith-based donors were committed to more distal outcomes such as changes in life outcomes of families or on the outcomes of outside investment in the community. The model led to a discussion of reasonableness of expectations and, in the end, to expanded evaluation indicators that included some of the more distal outcomes, but also to a greater appreciation by stakeholders of the intermediate milestones on the way to their preferred outcomes.

Are You Ready to Evaluate Outcomes?

While it is understood that the evaluation focus of the program will shift over time, here are some handy decision rules to decide whether it is time to shift the evaluation focus toward an emphasis on program outcomes:

- **Sustainability:** Political and financial will exists to sustain the intervention while the evaluation is conducted.
- **Fidelity:** Actual intervention implementation matches intended implementation. Erratic implementation makes it difficult to know what “version” of the intervention was implemented and, therefore, which version produced the outcomes.
- **Stability:** Intervention is not likely to change during the evaluation. Changes to the intervention over time will confound understanding of which aspects of the intervention caused the outcomes.
- **Reach:** Intervention reaches a sufficiently large number of clients (sample size) to employ the proposed data analysis. For example, the number of clients needed may vary with the magnitude of the change expected in the variables of interest (i.e., effect size) and the power needed for statistical purposes.
- **Dosage:** Clients have sufficient exposure to the intervention to result in the intended outcomes. Interventions with limited client contact are less likely to result in measurable outcomes as compared to interventions that provide more in-depth intervention with clients.

Illustrating Evaluation Focus Decisions

Because the appropriate evaluation focus is case-specific, let's apply these focus issues to a few different evaluation scenarios for the CLPP program. Think about two scenarios and how evaluation focus might differ for each.

- **Scenario 1**

At the 1-year mark, a neighboring community would like to adopt your program but wonders, "What are we in for?" Here you might determine that questions of efficiency and implementation are central to the evaluation. You would likely conclude this is a realistic focus, given the stage of development and the intensity of the program. Questions about outcomes would be premature.

- **Scenario 2**

At the 5-year mark, the auditing branch of your government funder wants to know, "Did you spend our money well?" Clearly, this requires a much more comprehensive evaluation, and would entail consideration of efficiency, effectiveness, possibly implementation, and cost-effectiveness. It is not clear, without more discussion with the stakeholder, whether research studies to determine causal attribution are also implied. Is this a realistic focus? At year 5, probably yes. The program is a significant investment in resources and has been in existence for enough time to expect some more distal outcomes to have occurred.

Note that in either scenario, you must also consider questions of interest to key stakeholders who are not necessarily intended users of the results of the current evaluation. Here those were defined to be advocates, who are concerned that families not be blamed for lead poisoning in their children, and housing authority staff, who are concerned that amelioration include estimates of costs and identification of less costly methods of lead reduction in homes. By year 5, these look like reasonable questions to include in the evaluation focus. At year 1, stakeholders might need assurance that you care about their questions, even if you cannot address them with this early evaluation.

Defining the Specific Evaluation Questions

These focus criteria just discussed identify the components of the logic model that are to be included in the evaluation focus, i.e., these activities, but not these; these outcomes, but not these. At this point, you convert the components of your focus into specific questions, i.e., implementation, effectiveness, efficiency, and attribution. Were my activities implemented as planned? Did my intended outcomes occur? Were the outcomes due to my activities as opposed to something else? If the outcomes occurred at some but not all sites, what barriers existed at less successful locations and what factors were related to success? At what cost were my activities implemented and my outcomes achieved?

Deciding On the Evaluation Design

Besides determining the evaluation focus and specific evaluation questions, at this point you also need to determine the appropriate evaluation design. There are many types of evaluation designs. Of chief interest in choosing the evaluation design is whether you are being asked to monitor progress on outcomes or whether you are also asked to show “attribution”—that progress on outcomes is related to your program efforts. These “attribution” questions may more appropriately be viewed as “research” as opposed to “program evaluation” depending on the level of scrutiny with which they are being asked.

Three general types of research designs are commonly recognized: experimental, quasi-experimental, and non-experimental/observational. Traditional program evaluation typically uses the third type, but all three are presented here because, over the life of the program, traditional evaluation approaches may need to be supplemented with other studies that look more like research.

Experimental designs use random assignment to compare the outcome of an intervention on one or more groups with an equivalent group or groups that did not receive the intervention. For example, you could select a group of similar schools, and then randomly assign some schools to receive a prevention curriculum and other schools to serve as controls. 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 onset or risk behavior longer than students in the control schools, you could attribute the success to your program.

However, in community settings it is hard, or sometimes even unethical, to have a true control group. While there are some solutions that preserve the integrity of experimental design, 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 outcomes of environmental tobacco smoke (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.

Related to quasi-experimental design, comparing outcomes/outcome data among states and between one state and the nation as a whole are common and important ways to evaluate public health efforts. Such comparisons will help you establish meaningful benchmarks for progress. States can also compare their progress with that of states with a similar investment in their area of public health, or they can contrast their outcomes with the results that could be expected if their programs were similar to those of states with a larger investment.

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 public health 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, which can be highlighted with time-series analyses. For example, questions on many of the larger national 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, surveillance coordinators, and statisticians to make state and national comparisons an important component of your evaluation.

Observational designs are common in program evaluation. These include, but are not limited to, time-series analysis, 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 public health 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 a unique outcome is being assessed, 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 identifying barriers to and facilitators of change.

This issue of “causal attribution,” while often a central research question, may or may not need to supplement traditional program evaluation. The field of public health is under increasing pressure to demonstrate that programs are worthwhile, effective, and efficient. During the last two decades, knowledge and understanding about how to evaluate complex programs have increased significantly. Nevertheless, because programs are so complex, these traditional research designs described here may not be a good choice. As the World Health Organization notes, “the use of randomized control trials to evaluate health promotion initiatives is, in most cases, inappropriate, misleading, and unnecessarily expensive.”⁴²

Therefore, before choosing experimental or even quasi-experimental designs to supplement more traditional program 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. Even when there is desire or need to “prove” that the program was responsible for progress on outcomes, traditional research designs may not be the only or best alternative. Depending on how rigorous the proof needs to be, proximity in time between the implementation of the program and the progress on outcomes, or systematic elimination of other alternative explanations may be enough to persuade key stakeholders that the program is making a contribution. While these design alternatives often cost less and require less time, keep in mind that saving time and money should not be the

⁴² WHO European Working Group on Health Promotion Evaluation. *op cit.*

main criterion when selecting an evaluation design. It is important to choose a design that will measure what you need to measure and will meet both your immediate and long-term needs.

Another alternative to experimental and quasi-experimental models is a goal-based evaluation model, which uses predetermined program goals and the underlying program theory as the standards for evaluation, thus holding the program accountable to prior expectations. The CDC Framework’s emphasis on program description and the construction of a logic model sets the stage for strong goal-based evaluations of programs. 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 the measurement activities.

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.

Standards for Step 3 Focus the Evaluation Design

Standard	Questions
Utility	<ul style="list-style-type: none"> • What is the purpose of the evaluation? • Who will use the evaluation results and how will they use them? • What are special needs of any other stakeholders that must be addressed?
Feasibility	<ul style="list-style-type: none"> • What is the program's stage of development? • How intense is the program? • How measurable are the components in the proposed focus?
Propriety	<ul style="list-style-type: none"> • Will the focus and design adequately detect any unintended consequences? • Will the focus and design include examination of the experience of those who are affected by the program?
Accuracy	<ul style="list-style-type: none"> • Is the focus broad enough to detect success or failure of the program? • Is the design the right one to respond to the questions—such as attribution—that are being asked by stakeholders?

Checklist for Focusing the Evaluation Design

- Define the purpose(s) and user(s) of your evaluation.
- Identify the use(s) of the evaluation results.
- Consider stage of development, program intensity, and logistics and resources.
- Determine the components of your logic model that should be part of the focus given these “utility and “feasibility” considerations.
- Formulate the evaluation questions to be asked of the program components in your focus, i.e., implementation, effectiveness, efficiency, and attribution questions.
- Review evaluation questions with stakeholders, program managers, and program staff.
- Review options for the evaluation design, making sure that the design *fits* the evaluation questions.

Worksheet 3A
Focusing the Evaluation in the Logic Model

#	If this is the situation ...	Then these are the parts of the logic model, I would include in my evaluation focus:
1	Who is asking evaluation questions of the program?	
2	Who will use the evaluation results and for what purpose?	
3	In Step 1, did we identify interests of other stakeholders that we must take into account?	

Worksheet 3B
“Reality Checking” the Evaluation Focus

#	If this is my answer to these questions...	Then I would conclude the questions in my evaluation focus are/are not reasonable ones to ask right now.
1	How long has the intervention been underway?	
2	How intensive/ambitious is the intervention? Multi-faceted effort or simple intervention?	
3	How much resources (time and money) are able to be devoted to evaluation of this effort?	

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 3: Focus the Evaluation

As discussed earlier, the “right” program evaluation is case-specific. Any given evaluation may focus on either activities (process evaluation) or outcomes, or both. Indeed, most evaluations will include some emphasis on both process and outcome measures.

Process Evaluations

Process evaluations focus on the program’s activities and are used to assess whether a program is being implemented as intended. Process indicators for appropriate antibiotic use activities might include:

- Number and type of organizations participating in a coalition
- Number and type of coalition or committee meetings
- Number of consumers receiving educational materials
- Economic status and racial/ethnic background of people participating in educational activities
- Number of providers participating in educational activities (lectures, grand rounds, CME events, etc.)
- Number of providers receiving prescribing feedback
- Inclusion of appropriate antibiotic use messages in consumer and provider publications
- Inclusion of appropriate antibiotic use content in medical school curricula
- Staffing for services and programs
- Amount of money used for the program
- Amount of in-kind services donated to the program
- Number of advertisements placed in print, television, radio, outdoor media

Outcome Evaluations

Outcome evaluations are used to assess the degree to which program activities are making progress on the sequence of outcomes or effects that the program aims to address – in this case, appropriate antibiotic use and antibiotic resistance. Outcomes can be differentiated into short-term, intermediate, and long-term outcomes.

Short-term outcomes for appropriate antibiotic use programs might include:

- Increased public knowledge and awareness of appropriate antibiotic use messages
- Increased knowledge and awareness among providers of appropriate antibiotic use
- Increased skills among providers to communicate appropriate antibiotic use messages to patients

- Adoption of appropriate prescribing guidelines by provider practice or organization
- Changes in childcare or workplace policies regarding return to school or work after an illness

Intermediate outcomes for appropriate antibiotic use programs might include:

- Changes in norms among providers to favor appropriate antibiotic prescribing
- Increased adherence to appropriate antibiotic use prescribing guidelines among providers
- Decreased patient demand for antibiotics
- Increased adherence to prescribed antibiotics among consumers

Long-term outcomes for appropriate antibiotic use programs might include:

- Decreased inappropriate antibiotic prescribing
- Decreased inappropriate antibiotic use
- Decrease in the rate of rise of antibiotic resistance

Determining the Evaluation Focus

As discussed previously, there is no “right” evaluation focus for a program – the focus will differ with each evaluation situation. When selecting measures for evaluation, program staff and stakeholders may choose to focus on some or all components of the program, from program inputs to the most long-term intended results.

The following questions can help your evaluation team determine the appropriate focus for your evaluation:

- What is the purpose of the evaluation?
- Who will use the evaluation results and how?
- How long has the program been in operation?
- How intensive is the program?

For example, stakeholders with a new program may decide that the primary purpose of their evaluation is to understand how doctors and other healthcare providers use educational materials with their patients. Measures of interest might include descriptions of materials currently used by providers and providers’ subjective ratings of these materials, as well as their appraisals of proposed materials. The results of the evaluation would then be used to develop appropriate materials and facilitate their use.

A program that has been in existence for a few years would want to measure its success in producing its intended outcomes, which might include changes in knowledge, attitudes, or behaviors. Even though this evaluation team may focus on outcome evaluation, they would also want to include a few process measures to give program staff information about whether the program has been implemented as planned and to describe the resources used to deliver the program.

While the overarching goal and ultimate outcome of appropriate antibiotic use programs is to slow the spread of resistance to antibiotics, there are several reasons why antibiotic resistance is almost never the only outcome studied in an evaluation of appropriate antibiotic use programs and often is not measured at all. First of all, all antibiotic use, whether inappropriate or appropriate, promotes resistance, and there are biological reasons why resistance rates may never decrease once they have increased.⁴³ Instead of looking toward decreases in resistance, programs might instead attempt to measure decreases in the rates of increase of antibiotic resistance. Furthermore, resistance rates are slow to change, and the effects of program activities may not be seen for several years. In addition, it is difficult and costly to measure resistance rates. Finally, because a number of other factors influence resistance rates, it is difficult to prove a causal link between program activities and corresponding rates in resistance. For example, the introduction of the pneumococcal vaccine in 2000 has resulted in a decline in pneumococcal disease among young children and appears to be preventing infections caused by drug-resistant strains.⁴⁴ Therefore, most program evaluations will focus on more proximal outcomes for their evaluation measures.

In comparison, intermediate outcomes, such as patient and provider behaviors (e.g., antibiotic use and prescribing), are much easier to measure, and changes in these outcomes will be visible more quickly. Because of the strong data linking antibiotic use and resistance, decreases in antibiotic prescribing or antibiotic use can be expected to have an effect on decreasing the spread of antibiotic resistance.

Most evaluation teams will look to even more proximal outcomes, such as changes in knowledge and attitudes related to antibiotic use or skills to communicate appropriate antibiotic messages. These short-term outcomes are even easier to measure than behavior changes, and they can be expected to change more quickly. Behavior change theories show that knowledge, attitudes, and skills are precursors to behavior change, so changes in these short-term outcomes can be expected to contribute to desired changes in antibiotic use and prescribing.

⁴³ Levin, BR: Minimizing potential resistance: a population dynamics view. *Clinical Infectious Diseases* 2001;33 (Suppl 3): S161-9.

⁴⁴ Whitney CG, Farley MM, Hadler J, et al.: Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *The New England Journal of Medicine* 2003;348(18):1737-46.

Step 4: Gather Credible Evidence

Now that you have developed a logic model, chosen an evaluation focus, and selected your evaluation questions, your next task is to gather the evidence. The gathering of evidence for an evaluation resembles the gathering of evidence for any research or data-oriented project, with a few exceptions noted below.

What's Involved in Gathering Evidence?

Evidence gathering must include consideration of each of the following:

- Indicators
- Sources of evidence/methods of data collection
- Quality
- Quantity
- Logistics

Developing Indicators

Because the components of our programs are often expressed in global or abstract terms, indicators are specific, observable, and measurable statements that help define exactly what we mean or are looking for. For example, the CLPP model includes global statements such as “Children receive medical treatment” or “Families adopt in-home techniques.” The medical treatment indicator might specify the type of medical treatment, the duration, or perhaps the adherence to the regimen. Likewise, the family indicator might indicate the in-home techniques or the intensity or duration of their adoption. For example, “Families with EBLL children clean all window sills and floors with the designated cleaning solution each week” or “Families serve leafy green vegetables at three or more meals per week.” *Outcome indicators* such as these indicators provide clearer definitions of the global statement and help guide the selection of data collection methods and the content of data collection instruments.

The activities in your focus may *also* include global statements such as “good coalition,” “culturally competent training,” and “appropriate quality patient care.” These activities would benefit from elaboration into indicators, often called “*process indicators*.” What does “good” mean, what does “quality” or “appropriate” mean?

Keep the following tips in mind when selecting your indicators:

- Indicators can be developed for activities (process indicators) and/or for outcomes (outcome indicators).⁴⁵
- There can be more than one indicator for each activity or outcome.

⁴⁵ Note that if you are developing your evaluation after completing an evaluation plan, you may already have developed process or outcome *objectives*. If the objectives were written to be specific, measurable, action-oriented, realistic, and time-bound (so-called “SMART” objectives), then they may serve as indicators as well.

- The indicator must be focused and must measure an important dimension of the activity or outcome.
- The indicator must be clear and specific in terms of what it will measure.
- The change measured by the indicator should represent progress toward implementing the activity or achieving the outcome.

Consider CDC’s immunization program, for example. The table below lists the components of the logic model that were included in our focus in Step 3. Then each of these components has been defined in one or more indicators.

Table 4.1
Provider Immunization Program:
Indicators for Program Component in Our Evaluation Focus

Program Component	Indicator(s)
Provider training	A series of 3 trainings will be conducted in all 4 regions of the state
Nurse educator LHD presentations	Nurse educators will make presentations to 10 largest local health departments (LHDs)
Physicians peer ed rounds	Physicians will host peer ed rounds at 10 largest hospitals
Providers attend trainings and rounds	Trainings will be well attended and reflect good mix of specialties and geographic representation
Providers receive and use tool kits	50%+ of providers who receive tool kit will report use of it (or “call to action” cards will be received from 25% of all providers receiving tool kit)
LHD nurses conduct private provider consults	Trained nurses in LHDs will conduct provider consults with largest provider practices in county
Provider KAB increases	Providers show increases in knowledge, attitudes, and beliefs (KAB) on selected key immunization items
Provider motivation increases	Provider intent to immunize increases

You may need to develop your own indicators or you may be able to draw on existing indicators developed by others. Some large CDC programs have developed indicator inventories that are tied to major activities and outcomes for the program. Advantages of these indicator inventories:

- They may have been pre-tested for “relevance” and accuracy.
- They define the best data sources for collecting the indicator.
- There are often many potential indicators for each activity or outcome, ensuring that at least one will be appropriate for your program.
- Because many programs are using the same indicator(s), you can compare performance across programs or even construct a national summary of performance.

Selecting Data Collection Methods and Sources

Now that you have determined the activities and outcomes you want to measure and the indicators you will use to measure progress on them, you need to select data collection methods and sources from which to gather information on your indicators.

A key decision is whether there are existing data sources—*secondary* data collection—to measure your indicators or whether you need to collect new data—*primary* data collection.

Depending on your evaluation questions and indicators, some secondary data sources may be appropriate data collection sources. Some existing data sources that often come into play in measuring outcomes of public health programs:

- Current Population Survey and other U.S. Census files
- Behavioral Risk Factor Surveillance System (BRFSS)
- Youth Risk Behavior Survey (YRBS)
- Pregnancy Risk Assessment Monitoring System (PRAMS)
- Cancer registries
- State vital statistics
- Various surveillance databases
- National Health Interview Survey (NHIS)

Before using secondary data sources, ensure that they meet your needs. Although large ongoing surveillance systems have the advantages of collecting data routinely and having existing resources and infrastructure, some of them (e.g., Current Population Survey [CPS]) have little flexibility with regard to the questions asked in the survey, making it nearly impossible to use these systems to collect the special data you may need for your evaluation. By contrast, other surveys such as BRFSS or PRAMS are more flexible. For example, you might be able to add program-specific questions, or you might expand the sample size for certain geographic areas or target populations, allowing for more accurate estimates in smaller populations.

The most common primary data collection methods also fall into several broad categories. Among the most common are:

- Surveys, including personal interviews, telephone, or instruments completed in person or received through the mail or e-mail
- Group discussions/focus groups
- Observation
- Document review, such as medical records, but also diaries, logs, minutes of meetings, etc.

Choosing the “right” method from the many secondary and primary data collection choices must consider both the *context* in which it is asked (How much money can be devoted to collection and measurement? How soon are results needed? Are there ethical considerations?) and the *content* of the question (Is it a sensitive issue? Is it about a behavior that is observable? Is it something the respondent is likely to know?).

Some methods yield qualitative data and some yield quantitative data. If the question involves an abstract concept or one where measurement is poor, using multiple methods is often helpful. Insights from stakeholder discussions in Step 1 and the clarity on purpose/user/use obtained in Step 3 will usually help direct the choice of sources and methods. For example, stakeholders may know which methods will work best with some intended respondents and/or have a strong bias toward quantitative or qualitative data collection that must be honored if the results are to be credible. More importantly, the purpose and use/user may dictate the need for valid, reliable data that will withstand close scrutiny or may allow for less rigorous data collection that can direct managers.

Each method comes with advantages and disadvantages depending on the context and content of the data collection (see Table 4.2).

Table 4.2
Advantages and Disadvantages of Various Survey Methods

Method	Advantages	Disadvantages
Personal interviews	<ul style="list-style-type: none"> • Least selection bias: can interview people without telephones—even homeless people. • Greatest response rate: people are most likely to agree to be surveyed when asked face to face. • Visual materials may be used. 	<ul style="list-style-type: none"> • Most costly: requires trained interviewers and travel time and costs. • Least anonymity: therefore, most likely that respondents will shade their responses toward what they believe is socially acceptable.
Telephone interviews	<ul style="list-style-type: none"> • Most rapid method. • Most potential to control the quality of the interview: interviewers remain in one place, so supervisors can oversee their work. • Easy to select telephone numbers at random. • Less expensive than personal interviews. • Better response rate than for mailed surveys. 	<ul style="list-style-type: none"> • Most selection bias: omits homeless people and people without telephones. • Less anonymity for respondents than for those completing instruments in private. • As with personal interviews, requires a trained interviewer.
Instruments to be completed by respondent	<ul style="list-style-type: none"> • Most anonymity: therefore, least bias toward socially acceptable responses. • Cost per respondent varies with response rate: the higher the response rate, the lower the cost per respondent. • Less selection bias than with telephone interviews. 	<ul style="list-style-type: none"> • Least control over quality of data. • Dependent on respondent's reading level. • Mailed instruments have lowest response rate. • Surveys using mailed instruments take the most time to complete because such instruments require time in the mail and time for respondent to complete.

The text box below lists possible sources of information for evaluations clustered in three broad categories: people, observations, and documents.

Some Sources of Data

Who might you survey or interview?

- Clients, program participants, nonparticipants
- Staff, program managers, administrators
- Partner agency staff
- General public
- Community leaders or key members of a community
- Funders
- Representatives of advocacy groups
- Elected officials, legislators, policymakers
- Local and state health officials

What might you observe?

- Meetings
- Special events or activities
- On the job performance
- Service encounters

Which documents might you analyze?

- Meeting minutes, administrative records
- Client medical records or other files
- Newsletters, press releases
- Strategic plans or work plans
- Registration, enrollment, or intake forms
- Previous evaluation reports
- Records held by funders or collaborators
- Web pages
- Graphs, maps, charts, photographs, videotapes

When choosing data collection methods and sources, select those that meet your project's needs. Try to avoid choosing a data method/source that may be familiar or popular but does not necessarily answer your questions. Keep in mind that budget issues alone should not drive your evaluation planning efforts.

The four evaluation standards can help you reduce the enormous number of data collection options to a more manageable number that best meet your data collection situation. Here is a checklist of issues — based on the evaluation standards — that will help you choose appropriately:

Utility

- Purpose and use of data collection: Do you seek a “point in time” determination of a behavior, or to examine the range and variety of experiences, or to tell an in-depth story?
- Users of data collection: Will some methods make the data more credible with skeptics or key users than others?

Feasibility

- Resources available: Which methods can you afford?
- Time: How long until the results are needed?
- Frequency: How often do you need the data?
- Your background: Are you trained in the method, or will you need help from an outside consultant?

Propriety

- Characteristics of the respondents: Will issues such as literacy or language make some methods preferable to others?
- Degree of intrusion to program/participants: Will the data collection method disrupt the program or be seen as intrusive by participants?
- Other ethical issues: Are there issues of confidentiality or safety of the respondent in seeking answers to questions on this issue?

Accuracy

- Nature of the issue: Is it about a behavior that is observable?
- Sensitivity of the issue: How open and honest will respondents be in responding to the questions on this issue?
- Respondent knowledge: Is it something the respondent is likely to know?

Using Multiple Methods and Mixed Methods

Sometimes a single method is not sufficient to accurately measure an activity or outcome because the thing being measured is complex and/or the data method/source does not yield data that are reliable or accurate enough. Employing multiple methods (sometimes called “triangulation”) helps increase the accuracy of the measurement and the certainty of your conclusions when the various methods yield similar results. Mixed data collection methods refers to gathering both quantitative and qualitative data. Mixed methods can be used sequentially, when one method is used to prepare for the use of another, or concurrently, when both methods are used in parallel. An example of sequential use of mixed methods is when focus groups (qualitative) are used to develop a survey instrument (quantitative), and then personal interviews (qualitative and quantitative) are conducted to investigate issues that arose during coding or interpretation of survey data. An example of concurrent use of mixed methods would be using focus groups or open-ended personal interviews to help affirm the response validity of a quantitative survey.

Different methods reveal different aspects of the program. Consider some interventions related to tobacco control:

- You might include a group assessment of a school-based tobacco control program to hear the group’s viewpoint, as well as individual student interviews to get a range of opinions.
- You might conduct a survey of all legislators in a state to gauge their interest in managed care support of cessation services and products, and you might also interview certain

legislators individually to question them in greater detail.

- You might conduct a focus group with community leaders to assess their attitudes regarding tobacco industry support of cultural and community activities. You might follow the focus group with individual structured or semi-structured interviews with the same participants.

When the outcomes under investigation are very abstract or no one quality data source exists, combining methods maximizes the strengths and minimizes the limitations of each method. Using multiple or mixed methods can increase the cross-checks on different subsets of findings and generate increased stakeholder confidence in the overall findings.

Illustrations from Cases

Consider the provider immunization education and the childhood lead poisoning examples. Table 4.3 presents data collection methods/sources for each of the indicators presented earlier for the provider immunization education program. Table 4.4 shows both the indicators and the data sources for key components of the CLPP effort presented earlier. Note in both cases that the methods/sources can vary widely and that in some cases multiple methods will be used and synthesized.

Table 4.3
Provider Immunization Education Program:
Data Collection Methods and Sources for Indicators

Indicator(s)	Data Collection Methods/Sources
A series of 3 trainings will be conducted in all 4 regions of the state	Training logs
Nurse educators will make presentations to 10 largest local health departments (LHDs)	Training logs
Physicians will host peer ed rounds at 10 largest hospitals	Training logs
Trainings will be well-attended and reflect good mix of specialties and geographic representation	Registration information
50%+ of providers who receive tool kit will report use of it (or "call to action" cards will be received from 25% of all providers receiving tool kit)	Survey of providers Analysis/count of call-to-action cards
Trained nurses in LHDs will conduct provider consults with largest provider practices in county	Survey of nurses, survey of providers, or training logs
Providers show increases in knowledge, attitudes, and beliefs (KAB) on selected key immunization items	Survey of providers, or focus groups, or intercepts
Provider intent to immunize increases	Survey of providers, or focus groups, or intercepts

Table 4.4
CLPP: Indicators and Data Collection Methods/Sources

Logic Model Element	Indicator(s)	Data Source(s) and Method(s)
Outreach	High-risk children and families in the district have been reached with relevant information	Logs of direct mail and health fair contacts Demographic algorithm Geographic Information System (GIS) algorithm
Screening	High-risk children have completed initial and follow-up screening	Logs and lab data
Environment assessment	Environments of all children over EBLL threshold have been assessed for lead poisoning	Logs of environmental health staff
Case management	All children over EBLL threshold have a case management plan including social, medical, and environmental components	Case file of EBLL child
Family training	Families of all children over EBLL threshold have received training on household behaviors to reduce EBLL	Logs of case managers Survey of families
“Leaded” houses referred	All houses of EBLL children with evidence of lead have been referred to housing authority	Logs and case files
“Leaded” houses cleaned	All referred houses have been cleaned up	Follow-up assessment by environmental health staff Logs of housing authority

Quality of Data

A quality evaluation produces data that are reliable, valid, and informative. An evaluation is reliable to the extent that it repeatedly produces the same results, and it is valid if it measures what it is intended to measure. The advantage of using existing data sources such as the BRFSS, YRBS, or PRAMS is that they have been pretested and designed to produce valid and reliable data. If you are designing your own evaluation tools, you should be aware of the factors that influence data quality:

- The design of the data collection instrument and how questions are worded
- The data collection procedures
- Training of data collectors
- The selection of data sources
- How the data are coded
- Data management
- Routine error checking as part of data quality control

A key way to enhance quality of primary data collection is through a pretest. The pretest need not be elaborate but should be extensive enough to determine issues of logistics of data collection or intelligibility of instruments prior to rollout. Obtaining quality data involves trade-offs (i.e., breadth vs. depth). Thus, you and stakeholders must decide at the beginning of the evaluation process what level of quality is necessary to meet stakeholders' standards for accuracy and credibility.

Quantity of Data

You will also need to determine the amount of data you want to collect during the evaluation. There are cases where you will need data of the highest validity and reliability, especially when traditional program evaluation is being supplemented with research studies. But there are other instances where the insights from a few cases or a convenience sample may be appropriate. If you use secondary data sources, many issues related to quality of data—such as sample size—have already been determined. If you are designing your own data collection tool and your examination of your program includes research as well as evaluation questions, the quantity of data you need to collect (i.e., sample sizes) will vary with the level of detail and the types of comparisons you hope to make. You will also need to determine the jurisdictional level for which you are gathering the data (e.g., state, county, region, congressional district). Counties often appreciate and want county-level estimates; however, this usually means larger sample sizes and more expense. Finally, consider the size of the change you are trying to detect. In general, detecting small amounts of change requires larger sample sizes. For example, detecting a 5% increase would require a larger sample size than detecting a 10% increase. You may need the help of a statistician to determine adequate sample size.

Logistics and Protocols

Logistics are the methods, timing, and physical infrastructure for gathering and handling evidence. People and organizations have cultural preferences that dictate acceptable ways of asking questions and collecting information, and influence who is perceived as an appropriate person to ask the questions (i.e., someone known within the community versus a stranger from a local health agency). The techniques used to gather evidence in an evaluation must be in keeping with a given community's cultural norms. Data collection procedures should also protect confidentiality.

In outlining procedures for collecting the evaluation data, consider these issues:

- When will you collect the data? You will need to determine when (and at what intervals) it is most appropriate to collect the information. If you are measuring whether your objectives have been met, your objectives will provide guidance as to when to collect certain data. If you are evaluating specific program interventions, you might want to obtain information from participants before they begin the program, upon completion of the program, and several months after the program. If you are assessing the effects of a community campaign, you might want to assess community knowledge, attitudes, and behaviors among your target audience before and after the campaign.

- Who will be considered a participant in the evaluation? Are you targeting a relatively specific group (African-American young people), or are you assessing trends among a more general population (all women of childbearing age)?
- Are you going to collect data from all participants or a sample? Some programs are community-based, and surveying a sample of the population participating in such programs is appropriate. However, if you have a small number of participants (such as students exposed to a curriculum in two schools), you may want to survey all participants.
- Who will collect the information? Are those collecting the data trained and trained consistently? Will the data collectors uniformly gather and record information? Your data collectors will need to be trained to ensure that they all collect information in the same way and without introducing bias. Preferably, interviewers should be trained together and by the same person.
- How will the security and confidentiality of the information be maintained? It is important to ensure the privacy and confidentiality of the evaluation participants. You can do this by collecting information anonymously and making sure you keep data stored in a locked and secure place.
- If your examination of your program includes research as well as evaluation studies: Do you need approval from an institutional review board (IRB) before collecting the data? What will be your informed consent procedures?

You may already have answered some of these questions while selecting your data sources and methods.

Agreements: Affirming Roles and Responsibilities

Agreements summarize the evaluation procedures, clarify everyone's role and responsibilities, and describe how the evaluation procedures will be implemented. Elements of an agreement include statements concerning the intended users, uses, purpose, questions, design, and methods, as well as a summary of the deliverables, timeline, and budget. An agreement might be a legal contract, a memorandum of understanding, or a detailed protocol. Creating an agreement establishes a mutual understanding of the activities associated with the evaluation. It also provides a basis for modification if necessary.

Standards for Step 4: Gather Credible Evidence

Standard	Questions
Utility	<ul style="list-style-type: none"> • Have key stakeholders been consulted who can assist with access to respondents? • Are methods and sources appropriate to the intended purpose and use of the data? • Have key stakeholders been consulted to ensure there are no preferences for or obstacles to selected methods or sources? • Are there specific methods or sources that will enhance the credibility of the data with key user and stakeholders?
Feasibility	<ul style="list-style-type: none"> • Can the data methods and sources be implemented within the time and budget for the project? • Does the evaluation team have the expertise to implement the chosen methods? • Are the methods and sources consistent with the culture and characteristics of the respondents, such as language and literacy level? • Are logistics and protocols realistic given the time and resources that can be devoted to data collection?
Propriety	<ul style="list-style-type: none"> • Will data collection be unduly disruptive? • Are there issues of safety of respondents or confidentiality that must be addressed? • Are the methods and sources appropriate to the culture and characteristics of the respondents—will they understand what they are being asked?
Accuracy	<ul style="list-style-type: none"> • Are appropriate QA procedures in place to ensure quality of data collection? • Are enough data being collected,—i.e., to support chosen confidence levels or statistical power? • Are methods and sources consistent with the nature of the problem, the sensitivity of the issue, and the knowledge level of the respondents?

Checklist for Gathering Credible Evidence

- Identify indicators for activities and outcomes in the evaluation focus.
- Determine whether existing indicators will suffice or whether new ones must be developed.
- Consider the range of data sources and choose the most appropriate one.
- Consider the range of data collection methods and choose those best suited to your context and content.
- Pilot test new instruments to identify and/or control sources of error.
- Consider a mixed-method approach to data collection.
- Consider quality and quantity issues in data collection.
- Develop a detailed protocol for data collection.

Worksheet 4A
Evaluation Questions, Indicators, and Data Collection Methods/Sources

Logic Model Components in Evaluation Focus		Indicator(s) or Evaluation Questions	Data Method(s)/Source(s)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Worksheet 4B Data Collection Logistics

	Data Collection Method/Source	From whom will these data be collected	By whom will these data be collected and when	Security or confidentiality steps
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 4: Gather Credible Evidence

The stakeholder discussions in Step 1 and the program description in Step 2 led to the selection of an evaluation focus in Step 3. At this point, you have a set of program components – activities and outcomes – that will be used in the evaluation. Next, you will need to develop tangible indicators (evaluation measures) for these components and identify data sources for each of the measures. The following table lists examples of indicators for selected appropriate antibiotic use activities and outcomes, as well as some associated data sources (Table 4.5).

Table 4.5: Appropriate Antibiotic Use Programs: Indicators and Data

Activities	Indicators	Data Sources
Formation of state or local coalition to develop and implement appropriate antibiotic use efforts	<ul style="list-style-type: none"> • Number of coalition meetings • Number and type of organizations involved in coalition 	Sign-in sheets and meeting minutes
Implementation of media campaign	<ul style="list-style-type: none"> • Number of impressions for print, television, radio, and outdoor media ads 	Media tracking reports
Development of health education materials	<ul style="list-style-type: none"> • Number and type of materials 	Program logs
Outcomes	Indicators	Data Sources
Increased public knowledge and awareness of appropriate antibiotic use messages	<ul style="list-style-type: none"> • Percentage of people who believe antibiotics will not help cure colds and flus • Percentage of people who recall the content of appropriate antibiotic use media campaign 	Consumer surveys
Increased knowledge and awareness among providers of appropriate antibiotic use messages	<ul style="list-style-type: none"> • Percentage of providers who believe inappropriate prescribing contributes to antibiotic resistance • Percentage of providers who recall the content of appropriate antibiotic use media campaign 	Provider surveys
Improved skills among providers to communicate appropriate antibiotic use messages to consumers	<ul style="list-style-type: none"> • Percentage of providers who report talking to patients about when antibiotics work and when they do not work • Percentage of patients who report satisfaction with their provider's communication 	Provider surveys Patient satisfaction surveys
Increased social norms favoring appropriate antibiotic prescribing	<ul style="list-style-type: none"> • Percentage of providers who believe that their peers follow prescribing guidelines 	Provider surveys

Increased adherence to appropriate antibiotic use guidelines	<ul style="list-style-type: none"> Percentage of providers who indicate that they follow appropriate antibiotic use guidelines (e.g., providers use rapid antigen test or throat culture to diagnose streptococcal pharyngitis) 	Provider surveys Chart reviews
Decreased patient demand for antibiotics	<ul style="list-style-type: none"> Percentage of consumers who state they do not ask providers for antibiotics Percentage of providers who state that their patients do not demand antibiotics 	Consumer surveys Provider surveys
Increased adherence to prescribed antibiotics among consumers	<ul style="list-style-type: none"> Percentage of consumers who state they finish the course of antibiotics Percentage of consumers who report they do not share antibiotics with others 	Consumer surveys
Incorporation of prescribing guidelines by provider practices or organizations	<ul style="list-style-type: none"> Number of provider practices or organizations that adopt appropriate prescribing guidelines as policy 	Surveys or interviews with practices or organizations
Changes in childcare or workplace policies supportive of appropriate antibiotic use	<ul style="list-style-type: none"> Number of childcare centers or work sites that do not require use of antibiotics before returning after an illness 	Surveys or interviews with childcare centers or work site staffs
Decreased inappropriate antibiotic use	<ul style="list-style-type: none"> Rates of antibiotic use for non-specific upper respiratory illnesses Rates of children tested for group A strep before receiving antibiotics for sore throats 	Pharmacy data Health plan data Health Plan Employer Data and Information Set (HEDIS®) performance measures

Secondary Data Sources

In some cases, data to evaluate the effectiveness of appropriate antibiotic use programs can be found in existing data sources. Three key secondary data sources are described below.

- Health plan data** – Health plans can be an excellent source of population-based data on antibiotic prescribing and utilization. When data are combined from several health plans, it is possible to obtain a good representation of the entire population. In addition, for patients with pharmacy benefits, pharmacy dispensing can be captured and linked to visit data. However, there are several limitations of working with health plan data. Missing claims and misclassification of diagnoses are common. In addition, health plan data usually do not cover drugs not paid for by the plan (e.g., samples dispensed in the office or drugs paid for out-of-pocket). Furthermore, the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which protects the confidentiality of individually identifiable health information, may limit the ability of health plans to share these data unless all personal identifiers can be removed. While there may be significant limitations to using health plan data, this data remains one of the most precise and useful sources of information on antibiotic

prescribing. Coalitions that include health plans can not only explore the use of health plan data for evaluation, but they can also use this data as part of their interventions (e.g., providing prescribing feedback to providers or to support organizational changes).

- Pharmacy data – Several companies collect and process data from pharmaceutical records of a number of sources, including drug manufacturers, wholesalers, retailers, pharmacies, mail order, long-term care facilities, and hospitals. Both antibiotic prescribing data and antibiotic retail sales data can be purchased, and these data can be used to evaluate the impact of a program on antibiotic prescribing. Some systems allow for data to be broken down to the level of the individual provider, and this information can be shared with providers as part of an intervention to promote more appropriate prescribing. These data are primarily used by pharmaceutical companies, and costs may be prohibitive for appropriate antibiotic use programs.
- Medicaid data – Medicaid claims data have been used by some programs to assess changes in prescribing. These data are freely available and contain information on prescribing to Medicaid recipients. However, the same caveats apply as described above for health plan data regarding HIPAA regulations, difficulties in interpreting administrative data, and completeness of reporting. In addition, in some states, the privatization of Medicaid has made these data no longer centrally available.

Data Collection Tools

In many cases, programs will not be able to obtain the necessary data from secondary data sources and will need to collect their own data for evaluation. Rather than developing entirely new data collection tools, programs can often use or adapt parts of existing tools. Many state and local programs have developed surveys to assess the knowledge, attitudes, and behaviors of both consumers and providers related to antibiotic use and prescribing. CDC has collected a number of these evaluation tools and has facilitated discussions of the strengths and limitations of tools and specific questions. Check the CDC Get Smart website (<http://www.cdc.gov/getsmart>) for a list of campaign partners and their current activities and evaluation plans. You can contact local program coordinators directly or request assistance through CDC.

In addition, questions on appropriate antibiotic use have been included in the population-based surveys described below. Programs may be able to access state or local data from these surveys. Programs can also model questions after these when designing their own questionnaires.

- Behavioral Risk Factor Surveillance System (BRFSS) – The BRFSS is a telephone survey conducted by the health departments of all states, the District of Columbia, Puerto Rico, the Virgin Islands, and Guam with assistance from CDC. The BRFSS is the primary source of information for states and the nation on the health-related behaviors of adults and includes questions related to behaviors associated with preventable chronic diseases, injuries, and infectious diseases. States can add questions specific to their needs, and in recent years,

some states have added questions on appropriate antibiotic use. See <http://www.cdc.gov/brfss/index.htm> for more information.

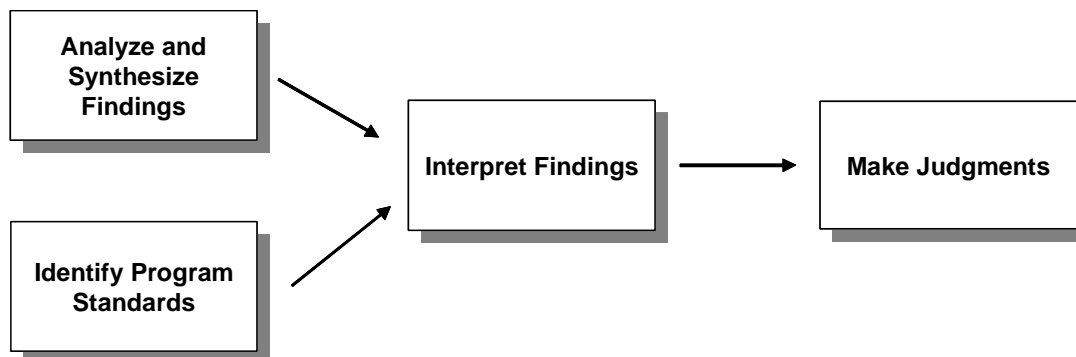
- FoodNet Population Survey – The Foodborne Diseases Active Surveillance Network (FoodNet) is the principal foodborne disease component of CDC's Emerging Infections Program (EIP). FoodNet conducts population-based telephone surveys to estimate the burden of acute diarrheal illness in the United States and the frequency of important exposures. The 2002-2003 FoodNet Population Survey included several questions to assess knowledge, attitudes, and behaviors surrounding appropriate antibiotic use. EIP sites may be able to use these data to document the need for their programs or to assess changes over time in knowledge, attitudes, and behaviors. Other states can model questions after these for local use and may be able to compare local results with those from FoodNet sites. See http://www.cdc.gov/foodnet/surveys/pop_cov.htm for more information.

Step 5: Justify Conclusions

Whether your evaluation is conducted to show program effectiveness, help improve the program, or demonstrate accountability, you will need to analyze and interpret the evidence gathered in Step 4. Step 5 encompasses analyzing the evidence, making claims about the program based on the analysis, and justifying the claims by comparing the evidence against stakeholder values.

Why is it Important to Justify Conclusions?

Why isn't this step called "analyze the data"? Because as central as data analysis is to evaluation, evaluators know that the evidence gathered for an evaluation does not necessarily speak for itself. As the figure below notes, conclusions become justified when analyzed and synthesized findings ("the evidence") are interpreted through the "prism" of values ("standards") that stakeholders bring, and then judged accordingly. Justification of conclusions is fundamental to utilization-focused evaluation. When agencies, communities, and other stakeholders agree that the conclusions are justified, they will be more inclined to use the evaluation results for program improvement.



The complicating factor, of course, is that different stakeholders may bring different and even contradictory standards and values to the table. As the old adage states, "where you stand depends on where you sit." Fortunately for those using the CDC Framework, the work of Step 5 benefits from the efforts of the previous steps: Differences in values and standards will have been identified during stakeholder engagement in Step 1. Those stakeholder perspectives will also have been reflected in the program description and evaluation focus.

Analyzing and Synthesizing the Findings

Data analysis is the process of organizing and classifying the information you have collected, tabulating it, summarizing it, comparing the results with other appropriate information, and presenting the results in an easily understandable manner. The five steps in data analysis and synthesis are straightforward:

- Enter the data into a database and check for errors. If you are using a surveillance system

such as BRFSS or PRAMS, the data have already been checked, entered, and tabulated by those conducting the survey. If you are collecting data with your own instrument, you will need to select the computer program you will use to enter and analyze the data, and determine who will enter, check, tabulate, and analyze the data.

- Tabulate the data. The data need to be tabulated to provide information (such as a number or %) for each indicator. Some basic calculations include determining
 - The number of participants
 - The number of participants achieving the desired outcome
 - The percentage of participants achieving the desired outcome.
- Analyze and stratify your data by various demographic variables of interest, such as participants' race, sex, age, income level, or geographic location.
- Make comparisons. When examination of your program includes research as well as evaluation studies, use statistical tests to show differences between comparison and intervention groups, between geographic areas, or between the pre-intervention and post-intervention status of the target population.
- Present your data in a clear and understandable form. To interpret your findings and make your recommendations, you must ensure that your results are easy to understand and clearly presented. Data can be presented in tables, bar charts, pie charts, line graphs, and maps.

In evaluations that use multiple methods, patterns in evidence are detected by isolating important findings (analysis) and combining different sources of information to reach a larger understanding (synthesis).

Setting Program Standards for Performance

“Program standards” as the term is used here—and not to be confused with the four evaluation standards discussed throughout this document—are the “benchmarks” that will be used to judge program performance. They reflect stakeholders’ values about the program and are fundamental to sound evaluation. The program and its stakeholders must articulate and negotiate the values that will be used to consider a program “successful,” “adequate,” or “unsuccessful.” Possible standards that might be used in determining these benchmarks:

- Needs of participants
- Community values, expectations, and norms
- Program mission and objectives
- Program protocols and procedures
- Performance by similar programs
- Performance by a control or comparison group
- Resource efficiency
- Mandates, policies, regulations, and laws
- Judgments of participants, experts, and funders
- Institutional goals
- Social equity

- Human rights.

When stakeholders disagree about standards/values, it may reflect differences in which outcomes are deemed most important. Or, stakeholders may agree on outcomes but disagree on the *amount* of progress on an outcome necessary to judge the program a success. This threshold for each indicator, sometimes called a “benchmark” or “performance indicator,” is often based on an expected change from a known baseline. For example, all CLPP stakeholders may agree that reduction in EBLL for program participants and provider participation in screening are key outcomes to judge the program a success. But, do they agree on how *much* of an EBLL decrease must be achieved for the program to be successful, or how *many* providers need to undertake screening of children for the program to be successful? In Step 5, you will negotiate consensus on these standards and compare your results with these performance indicators to justify your conclusions about the program. Performance indicators should be achievable but challenging, and should consider the program’s stage of development, the logic model, and the stakeholders’ expectations. Identifying and addressing differences in stakeholder values/standards early in the evaluation is helpful. If definition of performance standards is done *while* data are being collected or analyzed, the process can become acrimonious and adversarial.

Interpreting the Findings and Making Judgments

Judgments are statements about a program’s merit, worth, or significance. They are formed when findings are compared against one or more selected program standards. In forming judgments about a program:

- Multiple program standards can be applied.
- Stakeholders may reach different or even conflicting judgments.

Conflicting claims about a program’s quality, value, or importance often indicate that stakeholders are using different program standards or values in making their judgments. This type of disagreement can prompt stakeholders to clarify their values and reach consensus on how the program should be judged.

Tips To Remember When Interpreting Your Findings

- Interpret evaluation results with the goals of your program in mind.
- Keep your audience in mind when preparing the report. What do they need and want to know?
- Consider the limitations of the evaluation:
 - Possible biases
 - Validity of results
 - Reliability of results
- Are there alternative explanations for your results?
- How do your results compare with those of similar programs?
- Have the different data collection methods used to measure your progress shown similar results?
- Are your results consistent with theories supported by previous research?
- Are your results similar to what you expected? If not, why do you think they may be different?

Source: US Department of Health and Human Services. Introduction to program evaluation for comprehensive tobacco control programs. Atlanta, GA: US Department of Health and Human Services, Centers for Disease Control and Prevention, Office on Smoking and Health, November 2001.

Illustrations from Cases

Let's use the affordable housing program to illustrate the main points of this chapter about the sources of stakeholder disagreements and how they may influence an evaluation. For example, the various stakeholders may disagree about the key outcomes for success. Maybe the organization's staff, and even the family, deem the completion and sale of the house as most important. By contrast, the civic and community associations that sponsor houses and supply volunteers or the foundations that fund the organization's infrastructure may demand that home ownership produce improvement in life outcomes, such as better jobs or academic performance. Even when stakeholders agree on the outcomes, they may disagree about the amount of progress that needs to be made on these outcomes. For example, while churches may want to see improved life outcomes just for the individual families they sponsor, some foundations may be attracted to the program by the chance to change communities as a whole by changing the mix of renters and homeowners. As emphasized earlier in the chapter, it is important to identify these values and disagreements about values early in the evaluation so that consensus can be negotiated and so that program description and evaluation design and focus reflect the needs of the stakeholders who need and will use the data.

Standards for Step 5 Justify Conclusions

Standard	Questions
Utility	Have you carefully described the perspectives, procedures, and rationale used to interpret the findings? Have stakeholders considered different approaches for interpreting the findings?
Feasibility	Is the approach to analysis and interpretation appropriate to the level of expertise and resources?
Propriety	Have the standards and values of those less powerful or those most affected by the program been taken into account in determining standards for success?
Accuracy	Can you explicitly justify your conclusions? Are the conclusions fully understandable to stakeholders?

Checklist for Justifying Your Conclusions

- Analyze data using appropriate techniques.
- Check data for errors.
- Consider issues of context when interpreting data.
- Assess results against available literature and results of similar programs.
- If multiple methods have been employed, compare different methods for consistency in findings.
- Consider alternative explanations.
- Use existing standards (e.g., *Healthy People 2010* objectives) as a starting point for comparisons.
- Compare program outcomes with those of previous years.
- Compare actual with intended outcomes.
- Document potential biases.
- Examine the limitations of the evaluation.

Worksheet 5 Justify Conclusions

Question		Response
1	Who will analyze the data (and who will coordinate this effort)?	
2	How will data be analyzed and displayed?	
3	Against what “standards” will you compare your interpretations in forming your judgments?	
4	Who will be involved in making interpretations and judgments and what process will be employed?	
5	How will you deal with conflicting interpretations and judgments?	
6	Are your results similar to what you expected? If not, why do you think they are different?	
7	Are there alternative explanations for your results?	
8	How do your results compare with those of similar programs?	
9	What are the limitations of your data analysis and interpretation process (e.g., potential biases, generalizability of results, reliability, validity)?	
10	If you used multiple indicators to answer the same evaluation question, did you get similar results?	
11	Will others interpret the findings in an appropriate manner?	

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 5: Justify Conclusions

The next step of program evaluation is the analysis and interpretation of your findings. As discussed earlier, stakeholders can be involved in many aspects of program evaluation, and they are more likely to accept and use evaluation findings if they have been actively engaged in the evaluation from its inception. Furthermore, because stakeholders will define program success in different ways, it is important to understand the various viewpoints that stakeholders bring to an evaluation.

Step 6: Ensure Use of Evaluation Findings and Share Lessons Learned

The ultimate purpose of program evaluation is to use the information to improve programs. The purpose(s) you identified early in the evaluation process should guide the use of the evaluation results. The evaluation results can be used to demonstrate the effectiveness of your program, identify ways to improve your program, modify program planning, demonstrate accountability, and justify funding.

Additional uses include the following:

- To demonstrate to legislators or other stakeholders that resources are being well spent and that the program is effective.
- To aid in forming budgets and justify the allocation of resources.
- To compare outcomes with those of previous years.
- To compare actual outcomes with intended outcomes.
- To suggest realistic intended outcomes.
- To support annual and long-range planning.
- To focus attention on issues important to your program.
- To promote your program.
- To identify partners for collaborations.
- To enhance the image of your program.
- To retain or increase funding.
- To provide direction for program staff.
- To identify training and technical assistance needs.

What's involved in ensuring use and sharing lessons learned? Five elements are important in making sure that the findings from an evaluation are used:

- Recommendations
- Preparation
- Feedback
- Follow-up
- Dissemination

Making Recommendations

Recommendations are actions to consider as a result of an evaluation. Recommendations can strengthen an evaluation when they anticipate and react to what users want to know, and may undermine an evaluation's credibility if they are not supported by enough evidence, or are not in keeping with stakeholders' values.

Your recommendations will depend on the audience and the purpose of the evaluation (see text box). Remember, you identified many or all of these key audiences in Step 1, and have engaged many of them throughout as stakeholders. Hence, you have maximized the chances that the recommendations that you eventually make are relevant and useful to them. You know the information your stakeholders want and what is important to them. Their feedback early on in the evaluation makes their eventual support of your recommendations more likely.

Illustrations from Cases

Here are some examples, using the case illustrations, of recommendations tailored to different purposes and for different audiences:

Audience: Local provider immunization program.

Purpose of Evaluation: Improve program efforts.

Recommendation: Thirty-five percent of providers in Region 2 recalled the content of the monthly provider newsletter. To meet the current objective of a 50% recall rate among this population group, we recommend varying the media messages by specialty, and increasing the number of messages targeted through journals for the targeted specialties.

Audience: Legislators.

Purpose of Evaluation: Demonstrate effectiveness.

Recommendation: Last year, a targeted education and media campaign about the need for private provider participation in adult immunization was conducted across the state. Eighty percent of providers were reached by the campaign and reported a change in attitudes towards adult immunization—a twofold increase from the year before. We recommend the campaign be continued and expanded to include an emphasis on minimizing missed opportunities of providers to conduct adult immunizations.

Audience: County health commissioners.

Purpose of Evaluation: Demonstrate effectiveness of CLPP efforts.

Recommendation: In this past year, county staff identified all homes with EBLL children in targeted sections of the county. Data indicate that only 30% of these homes have been treated to eliminate the source of the lead poisoning. We recommend that you incorporate compliance checks for the lead ordinance into the county's housing inspection process and apply penalties for noncompliance by private landlords.

Audience: Foundation funding source for affordable housing program.

Purpose of Evaluation: Demonstrate fiscal accountability.

Recommendation: For the past 5 years, the program has worked through local coalitions, educational campaigns, and media efforts to increase engagement of volunteers and sponsors, and to match them with 300 needy families to build and sell a house. More than 90% of the

Some Potential Audiences for Recommendations

- Local programs
- The state health department
- City councils
- State legislators
- Schools
- Workplace owners
- Parents
- Police departments or enforcement agencies
- Healthcare providers
- Contractors
- Health insurance agencies
- Advocacy groups

families are still in their homes and making timely mortgage payments. But, while families report satisfaction with their new housing arrangement, we do not yet see evidence of changes in employment and school outcomes. We recommend continued support for the program but expansion to include an emphasis on tutoring and life coaching by the volunteers.

Preparation

Preparation refers to the steps taken to get ready to eventually use the evaluation findings. Through preparation, stakeholders can:

- Strengthen their ability to translate new knowledge into appropriate action.
- Discuss how potential findings might affect decision-making.
- Explore positive and negative implications of potential results and identify different options for program improvement.

Feedback

Feedback is the communication that occurs among everyone involved in the evaluation. Feedback, necessary at all stages of the evaluation process, creates an atmosphere of trust among stakeholders. Early in an evaluation, the process of giving and receiving feedback keeps an evaluation on track by keeping everyone informed about how the program is being implemented and how the evaluation is proceeding. As the evaluation progresses and preliminary results become available, feedback helps ensure that primary intended users and other stakeholders have opportunities to comment on evaluation decisions. Valuable feedback can be obtained by holding discussions and routinely sharing interim findings, provisional interpretations, and draft reports.

Follow-up

Although follow-up refers to the support that many users need throughout the evaluation process, in this step, in particular, it refers to the support that is needed after users receive evaluation results and begin to reach and justify their conclusions. Active follow-up can achieve the following:

- Remind users of the intended uses of what has been learned.
- Help to prevent misuse of results by ensuring that evidence is applied to the questions that were the evaluation's central focus.
- Prevent lessons learned from becoming lost or ignored in the process of making complex or political decisions.

Dissemination: Sharing the Results and the Lessons Learned From Evaluation

Dissemination is the process of communicating evaluation procedures or lessons learned to relevant audiences in a timely, unbiased, and consistent manner. Regardless of how

communications are structured, the goal for dissemination is to achieve full disclosure and impartial reporting. Planning effective communications requires

- Advance discussion of the reporting strategy with intended users and other stakeholders.
- Matching the timing, style, tone, message source, vehicle, and format of information products to the audience.

Some methods of getting the information to your audience include

- Mailings
- Websites
- Community forums
- Media (television, radio, newspaper)
- Personal contacts
- Listservs
- Organizational newsletters.

If a formal evaluation report is the chosen format, the evaluation report must clearly, succinctly, and impartially communicate all parts of the evaluation (see text box). The report should be written so that it is easy to understand. It need not be lengthy or technical. You should also consider oral presentations tailored to various audiences. An outline for a traditional evaluation report might look like this:

- **Executive Summary**
- **Background and Purpose**
 - Program background
 - Evaluation rationale
 - Stakeholder identification and engagement
 - Program description
 - Key evaluation questions/focus
- **Evaluation Methods**
 - Design
 - Sampling procedures
 - Measures or indicators
 - Data collection procedures
 - Data processing procedures
 - Analysis
 - Limitations
- **Results**
- **Discussion and Recommendations**

Tips for Writing Your Evaluation Report

- Tailor the report to your audience; you may need a different version of your report for each segment of your audience.
- Present clear and succinct results.
- Summarize the stakeholder roles and involvement.
- Explain the focus of the evaluation and its limitations.
- Summarize the evaluation plan and procedures.
- List the strengths and weaknesses of the evaluation.
- List the advantages and disadvantages of the recommendations.
- Verify that the report is unbiased and accurate.
- Remove technical jargon.
- Use examples, illustrations, graphics, and stories.
- Prepare and distribute reports on time.
- Distribute reports to as many stakeholders as possible.

Applying Standards

The three standards that most directly apply to Step 6—Ensure Use and Share Lessons Learned—are utility, propriety, and accuracy. As you use your own evaluation results, the questions presented in Table 6.1 can help you to clarify and achieve these standards.

Table 6.1
Standards for Step 6:
Ensure Use and Share Lessons Learned

Standard	Questions
Utility	<ul style="list-style-type: none"> • Do reports clearly describe the program, including its context, and the evaluation's purposes, procedures, and findings? • Have you shared significant mid-course findings and reports with users so that the findings can be used in a timely fashion? • Have you planned, conducted, and reported the evaluation in ways that encourage follow-through by stakeholders?
Feasibility	<ul style="list-style-type: none"> • Is the format appropriate to your resources and to the time and resources of the audience?
Propriety	<ul style="list-style-type: none"> • Have you ensured that the evaluation findings (including the limitations) are made accessible to everyone affected by the evaluation and others who have the right to receive the results?
Accuracy	<ul style="list-style-type: none"> • Have you tried to avoid the distortions that can be caused by personal feelings and other biases? • Do evaluation reports impartially and fairly reflect evaluation findings?

Evaluation is a practical tool that states can use to inform programs' efforts and assess their impact. Program evaluation should be well integrated into the day-to-day planning, implementation, and management of public health programs. Program evaluation complements CDC's operating principles for public health, which include using science as a basis for decision-making and action, expanding the quest for social equity, performing effectively as a service agency, and making efforts outcome-oriented. These principles highlight the need for programs to develop clear plans, inclusive partnerships, and feedback systems that support ongoing improvement. CDC is committed to providing additional tools and technical assistance to states and partners to build and enhance their capacity for evaluation.

Checklist for Ensuring that Evaluation Findings are Used and Sharing Lessons Learned

- Identify strategies to increase the likelihood that evaluation findings will be used.
- Identify strategies to reduce the likelihood that information will be misinterpreted.
- Provide continuous feedback to the program.
- Prepare stakeholders for the eventual use of evaluation findings.
- Identify training and technical assistance needs.
- Use evaluation findings to support annual and long-range planning.
- Use evaluation findings to promote your program.
- Use evaluation findings to enhance the public image of your program.
- Schedule follow-up meetings to facilitate the transfer of evaluation conclusions.
- Disseminate procedures used and lessons learned to stakeholders.
- Consider interim reports to key audiences.
- Tailor evaluation reports to audience(s.)
- Revisit the purpose(s) of the evaluation when preparing recommendations.
- Present clear and succinct findings in a timely manner.
- Avoid jargon when preparing or presenting information to stakeholders.
- Disseminate evaluation findings in several ways.

Worksheet 6A Communicating Results

	I need to communicate to this audience	This format would be most appropriate	This channel(s) would be most effective
1			
2			
3			
4			
5			
6			

Worksheet 6B Ensuring Follow-up

	The following will follow up with users of the evaluation findings	In this manner	This support is available for follow-up
1			
2			
3			
4			
5			
6			

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

Step 6: Ensure Use of Evaluation Findings and Share Lessons Learned

As discussed earlier, the way you use your evaluation findings and the recommendations you make will differ depending on your audience. Following are some hypothetical evaluation findings and potential recommendations that could be developed for the stakeholders listed.

Audience: State and local health department staff/health department administration.

Purpose of evaluation: Demonstrate effectiveness and expand program to new audiences.

Evaluation findings: Health education materials on appropriate antibiotic use were distributed to patients and providers at private practices in one major metropolitan area. Both consumers and providers reported high levels of knowledge and awareness of program messages following the intervention.

Recommendation: We recommend using the state and local health department infrastructure to expand distribution of materials to both patients and providers at public health clinics throughout the state.

Audience: Managed care organizations.

Purpose of evaluation: Demonstrate effectiveness; improve program efforts.

Evaluation findings: Appropriate prescribing guidelines were distributed to all providers in a managed care organization for the past two years, but prescribing rates did not change.

Recommendation: We recommend convening groups of providers and administrators from the managed care organization to discuss institutional barriers to changing prescribing practices (e.g., short visit times, formulary inventory) and suggestions for how to overcome these barriers (e.g., restructuring patient schedules, revisions to formulary).

Audience: Funding source.

Purpose of evaluation: Improve program efforts and reach multiple audiences.

Evaluation findings: For the past year, the state appropriate antibiotic use program has participated in an English-language media campaign that includes print, radio and television ads in conjunction with CDC's national media campaign. Half of the Caucasian population surveyed, but almost none of the American Indians, recalled the content of these ads.

Recommendation: To expand the reach of this campaign to include the state's large American Indian population, we recommend increased support to develop culturally and linguistically appropriate materials for this population.

Audience: Coalition members.

Purpose of evaluation: Develop and implement sustainability plan.

Evaluation findings: A coalition developed a presentation on antibiotic resistance and appropriate antibiotic use for community groups. Coalition members were trained to deliver the presentation and did so as their work schedules permitted. Within six months, the coalition had received over 100 requests for community presentations, and groups often had to wait several

months before a speaker was available. Participant evaluations of the community presentations were overwhelmingly positive and showed increases in participants' knowledge and awareness of appropriate antibiotic use messages following presentations.

Recommendation: To continue providing community presentations, the coalition will need to identify additional speakers. We recommend recruiting and training graduate students (e.g., public health, medical, or pharmacy students) to deliver presentations as part of their field work or community service requirements.

EVALUATING APPROPRIATE ANTIBIOTIC USE PROGRAMS

CASE STUDIES

The following case studies illustrate the use of CDC's framework in evaluating appropriate antibiotic use programs. These case studies provide concrete examples of the steps followed, problems encountered, and solutions found in planning and carrying out an evaluation and interpreting and sharing its results. Furthermore, the case studies illustrate the usefulness of the logic model as a tool for both program planning and program evaluation. Finally, the case studies reinforce the importance of building evaluation into program plans from the beginning.

The cases described are hypothetical, designed to include some typical intervention activities and demonstrate some of the common issues raised during program evaluation. These cases are not intended to serve as models or blueprints for program design or evaluation. Instead, it is our hope that the challenges raised by these case studies will help inspire solutions in the field.

Case Study: Clinic-Based Education for Patients and Providers

Background

A state health department received funding from the Centers for Disease Control and Prevention (CDC) to design and implement an appropriate antibiotic use campaign. This state is not part of CDC's Active Bacterial Core surveillance (ABCs), but the health department does have a sentinel surveillance system through which it collects and analyzes resistance data from several hospitals in the state. Epidemiologists, clinicians, and health educators within the health department identified antibiotic resistance and inappropriate antibiotic use as important problems in this state. Even though resistance rates decreased both nationally and within this state over the past several years, health department officials knew that the persistence of inappropriate antibiotic use could contribute to future increases in resistance rates. Doctors within the health department also reported that antibiotic prescribing rates remained high for viral upper respiratory conditions.

With CDC funding, the health department decided to hire a coordinator to organize a coalition of interested parties to develop and implement an intervention to promote more appropriate antibiotic use. Because CDC funds were limited, the group knew it would be important to evaluate this effort to ensure that resources were used wisely and to later advocate for continued funding.

Step 1: Engage Stakeholders.

Once the coordinator was hired, she began forming the coalition by identifying stakeholders to provide input on the development of the campaign and its evaluation. The program coordinator, who was housed within the communicable diseases branch of the health department, facilitated these early coalition meetings and invited health department staff with related interests and experience. Staff epidemiologists helped document the need for the program with surveillance data on local antibiotic resistance patterns. Other groups within the health department had considerable experience with community interventions and evaluations and thus were included in this effort. Public health nurses and health educators from maternal and child health talked about their experiences developing educational materials and campaigns for clinic use and described materials and approaches that had been effective.

Additional stakeholders for the evaluation of this program included physicians and nurse practitioners within the health department because program staff knew that any activities implemented within the health department clinics would need the support of the providers working in these clinics. As these clinicians joined the health department staff at coalition meetings and other planning meetings, they shared their ideas about what types of materials they would use in the health department clinics. They helped develop evaluation questions, and they helped program staff plan for the dissemination of evaluation results, particularly among healthcare providers. Physicians from the community were also identified and included through professional organizations (e.g., state medical associations), teaching hospitals, health plans, and

independent practice groups. Program staff made efforts to include local physicians who were leaders in the community because they would be able to help set norms favoring appropriate antibiotic prescribing.

The coalition also made efforts to engage patients or consumers – those who would be affected or served by the program – because they knew that by talking with patients, they would better understand the factors influencing antibiotic use, and they could also promote a sense of ownership for the intervention and its evaluation. The coalition considered forming a consumer advisory group but felt they did not have the time or resources. Instead they chose to informally talk with patients in health department clinic waiting rooms, and they used these discussions to elicit information about knowledge and behaviors surrounding antibiotic use. Through these talks, the program coordinator identified a few particularly interested and outspoken patients and invited them to attend regular coalition meetings and continue to be involved in planning and evaluation efforts.

Step 2: Describe the Program.

The stakeholder engagement proved influential in development of both the program and the evaluation. Coalition members initially had very different ideas about how to reduce inappropriate use of antibiotics. Providers cited patient demand as the primary reason they prescribed antibiotics when they might not be needed. Therefore, they felt that educating patients would lead to less demand and reduced antibiotic use. Consumers, however, were convinced that doctors and other providers were responsible for over-prescribing, citing short office visits and complicated explanations about their diagnosis and treatment plans. The program coordinator was able to guide the coalition through a planning process that helped members identify multiple factors influencing antibiotic prescribing and antibiotic use. Program staff knew from formative research that knowledge, attitudes, and behaviors of both consumers and providers contribute to inappropriate antibiotic use. They also knew from evaluations of other programs that efforts targeting both consumers and providers have proven to be most effective. In the end, the coalition decided to develop educational materials for both patients and providers and chose to distribute these materials through the health department's clinics in hopes of reaching a broad and diverse population. Because of the many references to lack of communication between patients and providers, they wanted to use the new materials to try to improve patient-provider interactions relating to antibiotic prescribing and use.

The coalition agreed that the overall goal of the program was to reduce inappropriate antibiotic use and decrease the spread of antibiotic resistance. Thanks to the extensive stakeholder engagement, the coalition also had a clear sense of the multiple paths they would need to follow. The resulting objectives and activities to meet this goal are summarized in the following table (Table 1).

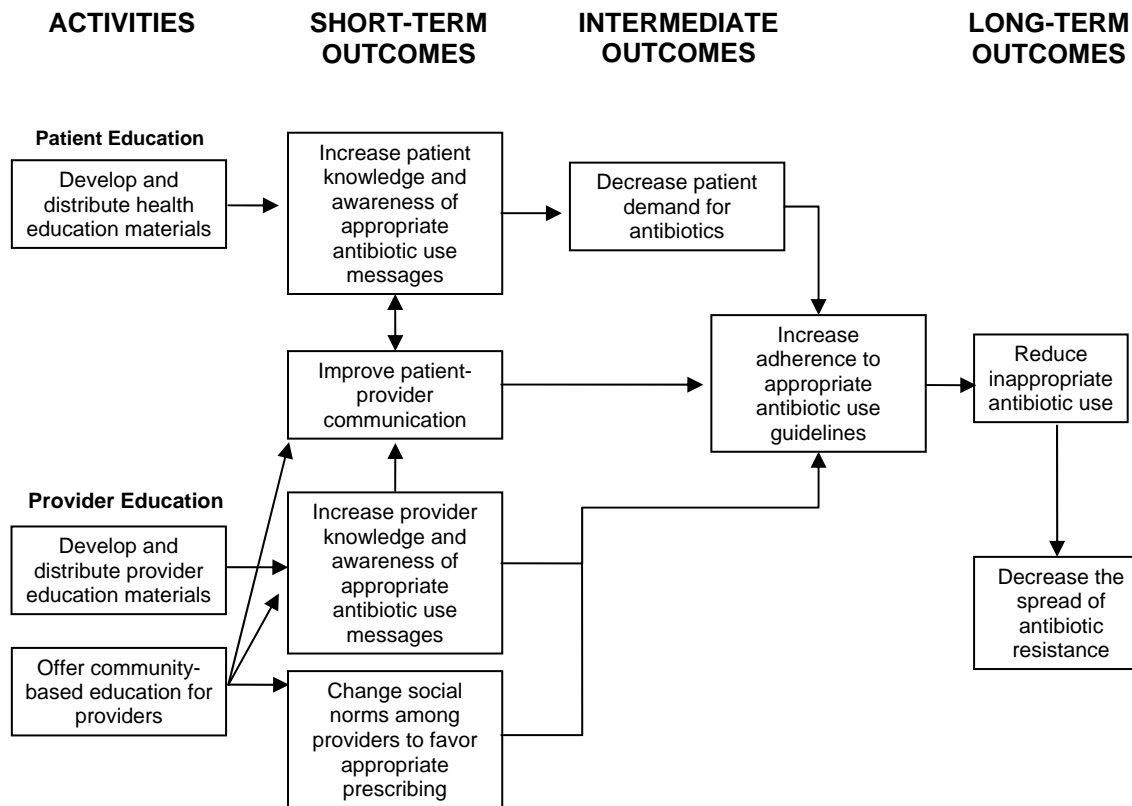
Table 1: Objectives and Activities

Objectives	Activities
Providers	
<ul style="list-style-type: none"> ▪ Increase adherence to appropriate antibiotic use guidelines. ▪ Increase provider knowledge and awareness of appropriate antibiotic use messages. ▪ Change social norms among providers to favor appropriate prescribing. 	<ul style="list-style-type: none"> ▪ Develop provider educational materials and distribute in health department clinics. ▪ Provide community-based professional education for providers.
Patients	
<ul style="list-style-type: none"> ▪ Decrease patient demand for antibiotics. ▪ Increase patient knowledge and awareness of appropriate antibiotic use messages. 	<ul style="list-style-type: none"> ▪ Develop patient education materials and distribute in health department clinics.
Patient-provider communication	
<ul style="list-style-type: none"> ▪ Improve patient-provider communication. 	<ul style="list-style-type: none"> ▪ Provide community-based professional education for providers. ▪ Develop educational materials for patients and providers and distribute in health department clinics.

For the patient education component, the coalition planned to develop and distribute health education materials (brochures, fact sheets, and posters) at the health department clinics. Educational materials were also developed for providers (detailing sheets modeled after those used by pharmaceutical companies), and these were mailed to physicians and nurse practitioners working at the health department clinics. Respected doctors in the community gave lectures at educational events for providers on topics such as appropriate antibiotic prescribing and tips for improving doctor-patient communication. Continuing medical education credits were provided for those attending.

These proposed program activities and their intended outcomes have been diagrammed in the following logic model. This model is a visual depiction of the activities and objectives listed above and shows the connection between specific activities and objectives (Exhibit 1).

Exhibit 1: Logic Model: Clinic-Based Education for Patients and Providers



Step 3: Focus the Evaluation Design.

During the stakeholder engagement (Step 1), the coalition had found that different stakeholders had very different priority outcomes for the project and therefore had very different ideas on where to focus the evaluation of the project. The group knew from early discussions with patients that they wanted to know if physicians would spend more time explaining whether or not they needed antibiotics and what they could do to feel better. The coalition had heard from doctors and nurse practitioners that they wanted to know if patients would ask for antibiotics less often and follow their prescriptions when they did receive antibiotics. Health department staff wanted to know if the intervention had any effect on antibiotic prescribing.

Since this evaluation was planned during the first year of the project implementation, the program staff decided that the overall purpose of the evaluation was to improve the program materials and strategies to increase the likelihood of reaching the program's intended outcomes. In particular, staff wanted to know if patients read and understood the materials because they wanted to make changes to the materials if needed. Similarly, staff wanted to know if providers found the provider materials useful. They planned to make changes to the content and/or delivery of these materials if needed. Staff also wanted to know what providers thought of the

community-based education and whether participation in these activities had any effect on communication skills or social norms.

Even at this early stage, program staff and other stakeholders expected to see some outcomes achieved. As mentioned earlier, stakeholders were interested in different sets of outcomes. Providers hoped to see decreased demand for antibiotics and increased adherence to prescribed antibiotics. Patients hoped for improved communication with their providers. Health department staff thought that the intervention would help to improve communication between patients and providers, resulting in decreased patient demand and greater patient satisfaction. Stakeholders decided to measure patient-provider communication because both patients and health department staff had explicitly named communication as an outcome of interest. In addition, improved communication was expected to contribute to decreased patient demand for antibiotics, which was the primary outcome of interest for providers.

Program staff chose not to measure antibiotic prescribing rates in the first year. Even though there was a high level of interest in this long-term outcome among the health department staff and clinic providers, the coalition decided to focus their outcome evaluation efforts on patient-provider communication (a short-term outcome) since resources for this evaluation were limited. In addition, because the overall purpose of the evaluation was to improve program efforts, program staff also included several process measures to document the implementation of program activities and measures to evaluate satisfaction with the new materials.

Based on the discussion above, the following evaluation questions were developed:

- Were patient and provider educational materials developed and distributed as planned?
- Was the community-based education for providers developed and implemented as planned?
- How satisfied were patients and providers with the materials and community-based education?
- Did patient-provider communication improve as a result of the intervention?

Step 4: Gather Credible Evidence.

The coalition decided to use both qualitative and quantitative data to best understand the implementation and effects of the program. The table below summarizes their data collection plan (Table 2). To answer the first two evaluation questions, which focused on the implementation of the program, staff collected and reviewed program logs, registration forms, and sign-in sheets.

Program staff wanted to hear from a large number of patients to assess satisfaction levels with the new educational materials, so they designed a short questionnaire to be completed by patients after their visits. They also interviewed a small sample of patients following their visits to get more in-depth and qualitative information. Program staff developed a similar questionnaire for providers to assess their satisfaction with provider materials and lectures. Staff also interviewed a small sample of providers to add qualitative data about how providers used the materials.

Improving patient-provider communication was one of the project’s objectives. At first, program staff had hoped to use a pretest-posttest design to survey patients before and after introducing the educational materials. As the intervention planning progressed, staff realized they did not have time to develop the questionnaire and survey patients before the planned launch of the intervention. More importantly, they had no way of predicting when clinic patients would return to the clinic, so they would not be able to easily collect pre-intervention and post-intervention questionnaires from the same patients to then compare results and measure improvements in communication. Instead, program staff used a posttest-only design and included questions on the patient questionnaires and interviews to assess their understanding of their providers’ explanations of diagnosis and treatment.

Although program staff were not able to use a pretest-posttest design with patients, they were able to do so with providers. Staff talked about conducting observations of patient-provider interactions to measure provider communication skills, but they did not have the funding or staffing to do this. Instead, providers were surveyed and interviewed both before and after the introduction of the new materials to assess their ability to explain antibiotic use to their patients.

Table 2: Evaluation Questions, Indicators and Data Sources

Evaluation Questions	Indicators	Data Sources
Were patient and provider educational materials developed and distributed as planned?	Number of materials developed Number of materials distributed	Program logs
Was the community-based education for providers developed and implemented as planned?	Number of educational events held Number of providers attending events by medical specialty	Registration forms Sign-in sheets
How satisfied were patients and providers with the materials and community-based education?	Percentage of patients who report satisfaction with materials Percentage of providers who report satisfaction with materials Percentage of providers who report use of materials Numbers of materials distributed by providers to patients Percentage of providers who report satisfaction with community-based education	Patient questionnaires Patient interviews Provider questionnaires Provider interviews
Did patient-provider communication improve as a result of the intervention?	Percentage of patients who state that they understand providers’ explanation of diagnosis and treatment Percentage of providers who state they are able to explain antibiotic use to patients	Patient questionnaires Patient interviews Provider questionnaires Provider interviews

Step 5: Justify the Conclusions.

Research staff at the health department analyzed data from the patient and provider questionnaires and interviews. Providers reported high levels of satisfaction with the materials but much lower levels of satisfaction with the community-based education. Providers also reported high levels of use of the new materials. Qualitative data from the provider interviews showed that providers were very satisfied with the materials and felt better able to talk about antibiotic use with their patients after the intervention than they did before. Many of the providers said that the new materials were good “tools” for patient education and helped them focus their discussions with patients around antibiotic use.

Patients reported much lower levels of satisfaction with the materials than did the providers. Qualitative data from the interviews helped staff understand this dissatisfaction. Many patients felt the materials were overly complicated and did not provide clear explanations of what to do to relieve their symptoms when antibiotics were not necessary. Patients, like providers, positively rated patient-provider communication, with the majority of patients saying they had received clear explanations of their diagnosis and treatment from their providers. A significant proportion of patients said that although they understood their providers, they left their visits with some unanswered questions.

Program staff were not entirely sure how to interpret these results, and stakeholders met to discuss the relative importance of the various findings. Stakeholders were pleased with the findings from the questionnaires and interviews showing that both patients and providers reported fairly high levels of communication and that providers reported improved communication following the introduction of the new materials. They were also pleased that providers reported high levels of satisfaction with program materials. However, they were surprised to find that patients did not understand the materials, especially since most said that they understood their providers’ explanations. While the intervention appeared successful as judged by the outcome of patient-provider communication, the group wanted to address the apparent lack of patient satisfaction with the materials. Stakeholders also felt that the project could be improved since patients reported leaving visits with unanswered questions.

Step 6: Ensure Use and Share Lessons.

Based on the discussion among stakeholders regarding the interpretation of data and conclusions about program success, the following recommendations were developed:

Coalition members recommended that patient materials be revised in the future with significant input from patients in order to improve comprehension and satisfaction. Coalition members decided to discontinue the community-based education due to the relative lack of satisfaction.

Some of the coalition members raised the idea that nurses, medical assistants, and other clinic staff could be an untapped resource for health education. Patients had complained of not having enough time with providers, and providers thought that these other clinic staff could help provide the additional education that patients wanted. Health department staff decided to distribute the

provider materials to other clinic staff, hold a training to introduce them to the project and gain their support, and look for opportunities for these staff to take on roles in educating patients on appropriate antibiotic use.

Given the high levels of satisfaction with provider materials and high ratings for communication between patients and providers, the health department designated funds to continue the project with the changes noted above. Stakeholders decided to collect more data on the intended outcomes of the project (i.e., knowledge and awareness of appropriate antibiotic use messages and antibiotic prescribing) during the second year of implementation.

Case Study: Media Campaign

Background

Researchers and administrators at a large health plan monitored antibiotic prescribing for several years and found increases in the number of antibiotics prescribed, as well as dramatic increases in health plan expenditures for antibiotics. They knew from national studies that many antibiotics are prescribed unnecessarily for upper respiratory infections and wanted to reduce inappropriate prescribing to both improve patient care and cut health plan costs. The researchers and administrators brought in physicians and other providers at the health plan to discuss the problem of antibiotic overuse and to develop possible approaches. The providers all cited patients' lack of knowledge about proper antibiotic use and high expectations for antibiotics as the main factors contributing to over-prescribing. As a result, the health plan decided to develop a media campaign to try to change public knowledge and awareness about antibiotic use. Since their data showed particularly high antibiotic utilization among young children, they decided to focus on parents of this population.

A coalition composed of health plan staff, medical professional groups, healthcare providers, public health department staff, and university researchers was formed to develop the media campaign. The coalition hired a project coordinator with funds from the Centers for Disease Control and Prevention (CDC) and began planning the media campaign. The group decided to use the radio public service announcement (PSA) and the print poster developed for CDC's national campaign, Get Smart: Know When Antibiotics Work, since the timing of their campaign coincided with CDC's national media launch. During the first year, the coalition solicited and received funds from several of its member organizations and used these funds to print posters and distribute them in community settings (i.e., community pediatric clinics, libraries), and to promote placement of the radio PSAs. The campaign was launched in two communities with high levels of membership in the participating health plan. Staff collected data to document the implementation of the campaign (e.g., number of ads placed) but did not measure any outcomes of the campaign. At the end of the campaign's first year, the coalition decided to implement the media campaign in two new communities and to develop and implement a more comprehensive evaluation plan.

Step 1: Engage Stakeholders.

The project coordinator convened a series of meetings to begin planning the evaluation of the expanded media campaign. A diverse group of stakeholders was identified, including coalition members (described above) and representatives from the local radio stations that had run the PSAs during the first year of the campaign. The coalition also wanted to include members of the target audience, so they made presentations about the campaign at school and community functions in the new target communities and recruited parents of young children to participate in the planning of the evaluation.

Stakeholders had very different views of what should be measured by this second evaluation. Health educators shared their experiences evaluating other media campaigns and recommended an assessment of exposure to campaign messages. Radio station representatives were also interested in exposure and offered to collect measures of the reach and frequency of message exposure. The coalition members who had contributed funds towards the media campaign felt it was more important to show some “results” this year, stating that they needed to report back to their organizations and show that their money had been put to good use in order to keep their organizations engaged. These organizations were willing to accept a process evaluation in the first year, but they felt increased pressure to show results during the second year.

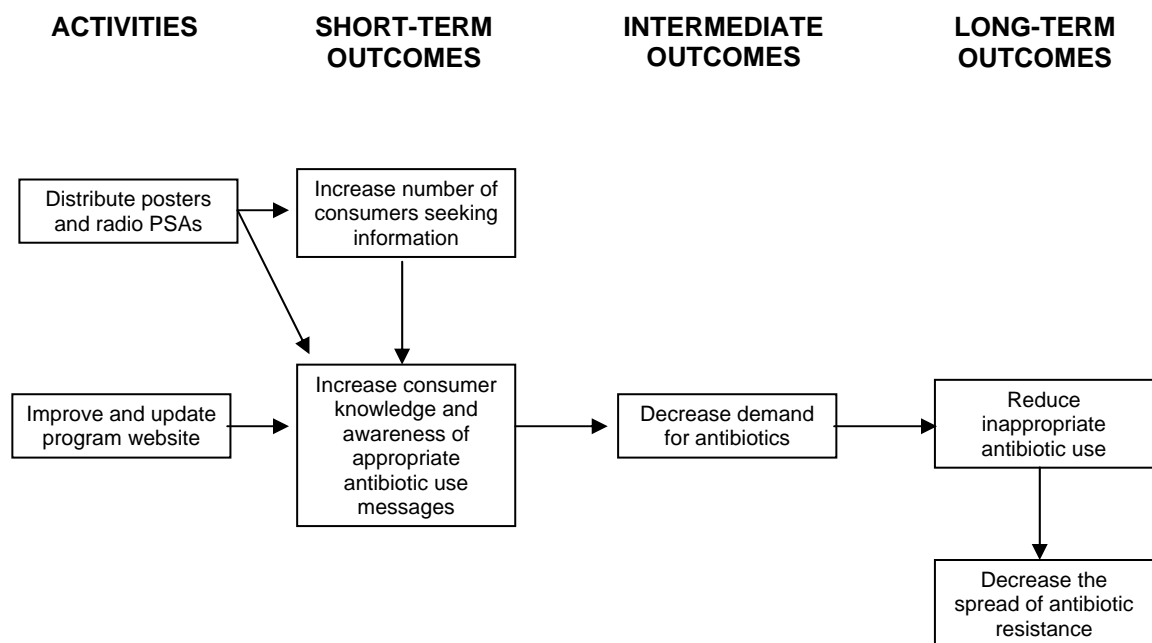
The project coordinator and other coalition members wanted to use the evaluation to document adherence to their implementation plans and to find out if consumers were motivated by the campaign to seek additional information about appropriate antibiotic use. The coalition had developed a website at the end of the first year, and the group now decided to update and improve the website and to include their Web address on all media pieces for the second year of the campaign. Finally, the parents involved contributed yet another perspective to the evaluation. They said they often felt overwhelmed by the quantity of ads and messages in the media, and they wanted to know if parents even noticed or paid attention to the campaign PSAs and posters.

Step 2: Describe the Program.

As a result of stakeholder discussions and the need to show “results,” the coalition realized they needed to make the goals and objectives of the media campaign more explicit and discuss what would constitute success for the project. The coalition had never formally stated goals or objectives for the project, but they now saw the evaluation of the second year of the campaign as a great opportunity to engage the coalition in this process.

The coalition agreed that the long-term goal of the program was to decrease the spread of antibiotic resistance. The project coordinator then helped the coalition draft a logic model to show the relationship between program activities (the media campaign) and desired outcomes leading up to the program’s long-term goal. Stakeholders identified intermediate milestones between the implementation of the project activities and this long-term goal, including: increased consumer knowledge and awareness of appropriate antibiotic use messages, increased number of consumers seeking information (i.e., from the program website), decreased patient demand for antibiotics, and reduced inappropriate antibiotic use. The following logic model visually depicts the activities and intended outcomes of the project and the hypothesized relationships among them (Exhibit 1).

Exhibit 1: Logic Model: Media Campaign



Drafting the logic model helped the group gain clarity regarding its goals and objectives. In the end, three of the outcomes from the logic model were adopted as the program objectives, as summarized in the following table (Table 1).

Table 1: Objectives and Activities

Objectives	Activities
<ul style="list-style-type: none"> ▪ Increase knowledge and awareness of appropriate antibiotic use messages. ▪ Increase number of consumers seeking information on antibiotic use. ▪ Decrease patient demand for antibiotics. 	<ul style="list-style-type: none"> ▪ Distribute posters and radio PSAs. ▪ Improve and update program website.

Step 3: Focus the Evaluation Design.

As noted, during the stakeholder engagement in Step 1, the coalition found that different stakeholders had differing views as to how to focus the evaluation of this program. The project coordinator and some of the other health department staff wanted to focus on implementation of the campaign and exposure to media messages, while some of the coalition members (particularly those who had contributed funding) wanted to measure results or outcomes. Parents were most interested in whether or not consumers saw the posters or heard the radio PSAs.

As part of the process evaluation conducted during year one of the campaign, coalition members had documented the implementation of the campaign and collected data on the number of media pieces placed, the timing and location of these placements, and the estimated number of audience impressions (or viewings) for each type of media used. Stakeholders decided to continue to

collect this process data to make sure that the media campaign activities were implemented as planned; however, they decided it was important to add outcome measures this year since the evaluation would be used to justify continued funding from their partners.

The coalition met to select primary outcomes of interest. Some of the physicians from participating health plans wanted to measure antibiotic use and suggested using the number of prescriptions as an indicator. Marketing experts on the coalition convinced the group that, based on much research and evaluation, it was unrealistic to expect significant behavior change as a result of a media campaign alone. Instead they suggested looking at shorter-term outcomes such as changes in knowledge and awareness that are likely to precede, yet eventually contribute to, the desired behavior change. Those representing the funding organizations lobbied hard to look beyond knowledge and awareness to some actual “results,” believing that they should see behavior change at this point in the campaign if it were working. In the end, the group chose to focus on both short-term and intermediate outcomes, including knowledge and awareness, information-seeking, and measures of patient demand for antibiotics. They believed that all of these factors could be influenced by a media campaign and that changes in patient demand would be seen as tangible results by the coalition members.

Based on the discussion above, the following evaluation questions were developed:

- Was the media campaign implemented as planned?
- Did consumer knowledge and awareness of appropriate antibiotic use messages increase?
- Did consumers who were exposed to the media campaign seek information about antibiotic use?
- Did patient demand for antibiotics decrease?

Step 4: Gather Credible Evidence.

The first of the group’s evaluation questions is a process question and looks at the implementation of program activities. To determine whether the campaign was implemented as planned, the program coordinator replicated the process evaluation employed in the first year. The coordinator collected samples of all the media materials in use and reviewed program logs documenting poster distribution. Coalition members from local radio stations facilitated the media tracking and provided figures for the numbers of PSAs aired and the estimated number of people who heard the ads in both of the new target communities.

The remaining evaluation questions look at outcomes, changes in things other than the program and its staff. The coalition decided that a survey was the best way to measure changes in knowledge and awareness of appropriate antibiotic use. They did not have funds to implement a survey of their own, but fortunately, one of the participating health plans agreed to add a few questions to an existing state-wide consumer telephone survey that was being used to assess community need for new pediatric clinics. This survey was being conducted with the same target population as the media campaign – parents of young children. A few questions were drafted on knowledge and awareness of appropriate antibiotic use messages, and these questions were added to the survey. Some parents of young children who had participated in early coalition meetings as stakeholders reviewed the questions to ensure comprehension.

Next, coalition members proposed using the number of hits to the program website as a measure of consumers seeking additional information about antibiotic use. Finally, the group discussed a variety of ways to measure patient demand for antibiotics. Coalition members wanted to objectively measure patient demand, but they did not have the resources to observe or record patient-provider encounters. They also considered surveying providers to find out whether or not their patients asked them for antibiotics, but they did not have funds to develop and implement this type of survey. Instead they decided to include a question on the consumer telephone survey asking parents if they had requested antibiotics for their children, and they used this self-reported data as the measure of patient demand.

Because of the prominence of the issue of antibiotic resistance and the ongoing CDC national media campaign, the coalition knew it would be difficult to isolate the unique contributions of their efforts to change consumer knowledge and demand. They considered adding a simple question to the survey to determine consumers' source of information on the topic but decided they would need a stronger case to keep the funding organizations engaged. Fortunately, the health plan survey was being administered in multiple communities, and the project staff was able to choose two communities with similar demographics to serve as control communities. The health plan survey was conducted in the two new campaign communities and two control communities both before and after the introduction of media messages into the new campaign communities. Within each community, phone calls were made to random households with young children. Half of the surveys were conducted before the media campaign began, and the other half were conducted afterwards with different respondents. Because of this design, changes in knowledge, awareness, and demand were measured at the population level, rather than the individual level. The table below summarizes their data collection plan (Table 2).

Table 2: Evaluation Questions, Indicators and Data Sources

Evaluation Questions	Indicators	Data Sources
Was the media campaign implemented as planned?	Number of posters and radio PSAs placed. Estimated number of people who see or hear ads.	Media materials, program logs, media tracking
Did consumer knowledge and awareness of appropriate antibiotic use messages increase?	Percentage of consumers who report seeing posters and hearing radio PSAs. Percentage of consumers who believe antibiotics are not useful for colds and flu. Percentage of consumers who are aware of the threat of antibiotic resistance.	Consumer telephone survey
Did consumers who were exposed to the media campaign seek information about antibiotic use?	Number of program website hits.	Website tracking
Did demand for antibiotics decrease?	Percentage of consumers who reported that their child had cold or flu symptoms and who also reported asking their provider for an antibiotic.	Consumer telephone survey

Step 5: Justify the Conclusions.

The data were collected and analyzed according to plans. The data from the process evaluation showed that all media materials were developed as planned but that ad placement varied dramatically between communities. One of the campaign communities documented two to three times the number of posters and ads placed and number of audience impressions as compared to the other campaign community.

Stakeholders met with the staff responsible for ad placement to better understand the differences in the number of ads placed and posters distributed in the various communities. Staff in both campaign communities had followed a protocol for contacting radio stations and seeking placement of the radio PSAs. The two campaign communities had received the same amount of funds for poster placement from the coalition and had distributed roughly the same number of posters with this funding. However, in one of the campaign communities, project staff had developed partnerships that resulted in increased public exposure to the campaign. Staff in this community had met with local hospitals and clinics and had succeeded in placing the radio PSAs on telephone recordings for callers on hold. In addition, the hospitals had provided in-kind donations of printing services and had distributed additional posters throughout their provider networks. Furthermore, these staff had worked closely with parents at community schools and had played the radio PSA and distributed campaign posters at school meetings within this community.

Program staff analyzed results from the consumer telephone survey and found that knowledge and awareness of appropriate antibiotic use messages increased in the community with high levels of message exposure as compared with the low-exposure community and the control communities. However, self-reported demand for antibiotics did not change significantly in any of the communities.

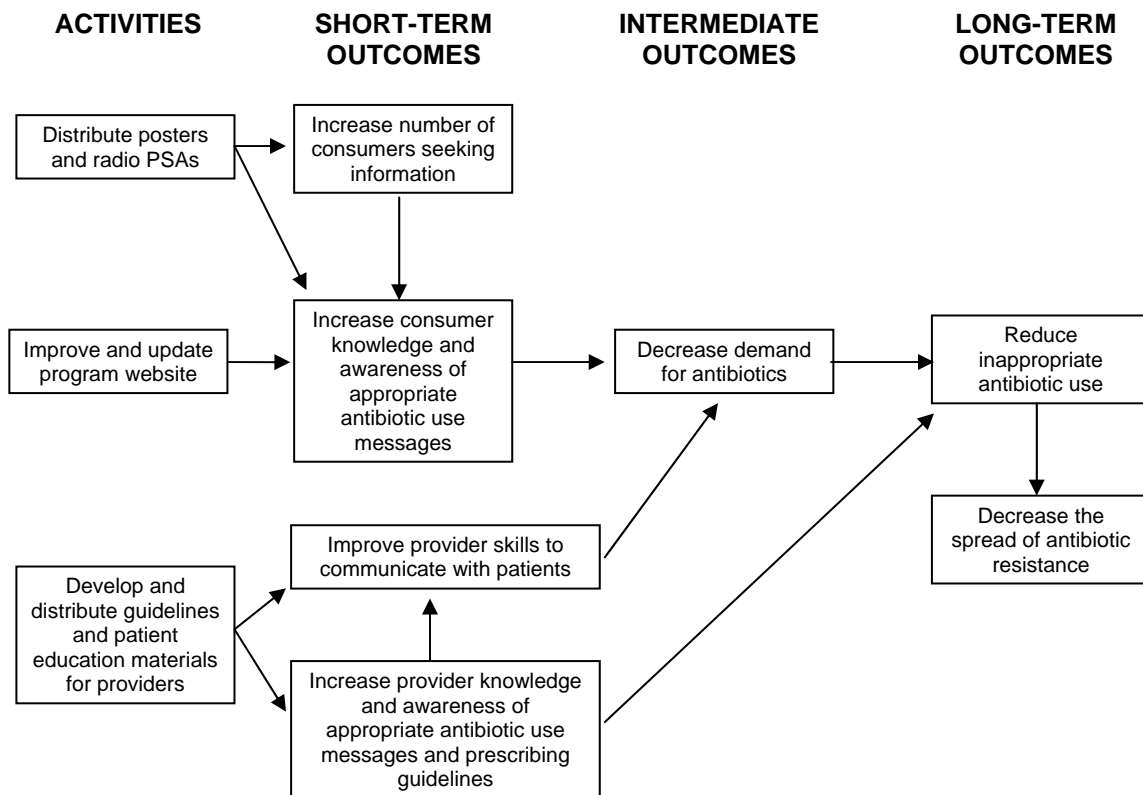
Website tracking showed a significant increase in the number of hits to the program website following the introduction of media messages, but the tracking software used did not allow the breakdown of totals by community. Stakeholders assumed the website hits came from the community with greater exposure to the campaign, but they were not able to document this. The group concluded that while providing a website as a resource for people seeking more information was an important component of the media campaign, the number of people seeking further information was not a useful evaluation indicator in this case given the limitations of their software.

Stakeholders were not surprised that demand for antibiotics did not change in the target community, which showed little or no improvement in knowledge and awareness of appropriate antibiotic use. With the exception of the marketing experts, stakeholders did, however, expect to see changes in demand as a result of increases in knowledge and awareness and were surprised to see no change in demand within the campaign community that documented increased knowledge and awareness. Coalition members concluded that the media campaign alone was not enough to lead to decreases in patient demand. Marketing experts on the coalition spoke of the need for a “supportive environment” to reinforce and supplement the knowledge and attitude changes in order for this behavior change to occur. The group hypothesized that in this case, healthcare

providers did not provide the supportive environment needed to result in decreased patient demand. Public knowledge and awareness had increased in one of the campaign communities, while provider knowledge and skills in this same community remained relatively unchanged. Even though healthcare providers had been exposed to campaign messages as part of the general media campaign, the campaign did not include specific activities or tools for providers. Stakeholders concluded that to achieve behavior change (i.e., decreased patient demand or decreased inappropriate antibiotic use), they would need to target healthcare providers in order to improve providers' knowledge and communication skills.

Following this stakeholder discussion, the project coordinator revised the logic model to include another pathway depicting provider education activities. The project coordinator proposed developing prescribing guidelines and patient education materials to distribute to healthcare providers. This activity was expected to contribute to increased provider knowledge and awareness of appropriate antibiotic use and to increased skills to communicate about appropriate antibiotic use with their patients. The following logic model shows the relationship between provider education and the hypothesized outcomes of these and other program activities (Exhibit 2).

Exhibit 2: Revised Logic Model: Media Campaign



Step 6: Ensure Use and Share Lessons.

Based on the discussion among stakeholders regarding the interpretation of data and conclusions about program success, the following recommendations were developed:

Stakeholders recommended continuing the media campaign and expanding the program to include a provider education component consisting of (at a minimum) the distribution of prescribing guidelines and patient education materials as tools to help improve patient-provider communication. Health plan partners suggested additional vehicles for provider communication and education within the health plan system, including internal newsletters and continuing medical education events. The stakeholder group also stressed the importance of including an evaluation component for any new provider education activities.

Because additional PSA placement and poster distribution had resulted in greater exposure and increased knowledge and awareness of campaign messages in two of the target communities, stakeholders were reminded of the importance of partnerships and recommended pursuing partnerships to expand the reach of future campaign efforts. Finally, stakeholders recommended that staff responsible for media placement in the various communities maintain frequent contact to share successful strategies and techniques.

Glossary

Accountability: The responsibility of program managers and staff to provide evidence to stakeholders and funding agencies that a program is effective and in conformance with its coverage, service, legal, and fiscal requirements.

Accuracy: The extent to which an evaluation is truthful or valid in what it says about a program, project, or material.

Activities: The actual events or actions that take place as a part of the program.

Attribution: The estimation of the extent to which any results observed are caused by a program, meaning that the program has produced incremental effects.

Breadth: The scope of the measurement's coverage.

Case study: A data collection method that involves in-depth studies of specific cases or projects within a program. The method itself is made up of one or more data collection methods (such as interviews and file review).

Causal inference: The logical process used to draw conclusions from evidence concerning what has been produced or “caused” by a program. To say that a program produced or caused a certain result means that, if the program had not been there (or if it had been there in a different form or degree), then the observed result (or level of result) would not have occurred.

Comparison group: A group not exposed to a program or treatment. Also referred to as a control group.

Comprehensiveness: Full breadth and depth of coverage on the evaluation issues of interest.

Conclusion validity: The ability to generalize the conclusions about an existing program to other places, times, or situations. Both internal and external validity issues must be addressed if such conclusions are to be reached.

Confidence level: A statement that the true value of a parameter for a population lies within a specified range of values with a certain level of probability.

Control group: In quasi-experimental designs, a group of subjects who receive all influences except the program in exactly the same fashion as the treatment group (the latter called, in some circumstances, the experimental or program group). Also referred to as a non-program group.

Cost-benefit analysis: An analysis that combines the benefits of a program with the costs of the program. The benefits and costs are transformed into monetary terms.

Cost-effectiveness analysis: An analysis that combines program costs and effects (impacts). However, the impacts do not have to be transformed into monetary benefits or costs.

Cross-sectional data: Data collected at one point in time from various entities.

Data collection method: The way facts about a program and its outcomes are amassed. Data collection methods often used in program evaluations include literature search, file review, natural observations, surveys, expert opinion, and case studies.

Depth: A measurement's degree of accuracy and detail.

Descriptive statistical analysis: Numbers and tabulations used to summarize and present quantitative information concisely.

Diffusion or imitation of treatment: Respondents in one group get the effect intended for the treatment (program) group. This is a threat to internal validity.

Direct analytic methods: Methods used to process data to provide evidence on the direct impacts or outcomes of a program.

Evaluation design: The logical model or conceptual framework used to arrive at conclusions about outcomes.

Evaluation plan: A written document describing the overall approach or design that will be used to guide an evaluation. It includes what will be done, how it will be done, who will do it, when it will be done, why the evaluation is being conducted, and how the findings will likely be used.

Evaluation strategy: The method used to gather evidence about one or more outcomes of a program. An evaluation strategy is made up of an evaluation design, a data collection method, and an analysis technique.

Ex ante cost-benefit or cost-effectiveness analysis: A cost-benefit or cost-effectiveness analysis that does not estimate the actual benefits and costs of a program but that uses hypothesized before-the-fact costs and benefits. This type of analysis is used for planning purposes rather than for evaluation.

Ex post cost-benefit or cost-effectiveness analysis: A cost-benefit or cost-effectiveness analysis that takes place after a program has been in operation for some time and that is used to assess actual costs and actual benefits.

Executive summary: A nontechnical summary statement designed to provide a quick overview of the full-length report on which it is based.

Experimental (or randomized) designs: Designs that try to ensure the initial equivalence of one or more control groups to a treatment group by administratively creating the groups through random assignment, thereby ensuring their mathematical equivalence. Examples of experimental or randomized designs are randomized block designs, Latin square designs, fractional designs, and the Solomon four-group.

Expert opinion: A data collection method that involves using the perceptions and knowledge of experts in functional areas as indicators of program outcome.

External validity: The ability to generalize conclusions about a program to future or different conditions. Threats to external validity include selection and program interaction, setting and program interaction, and history and program interaction.

File review: A data collection method involving a review of program files. There are usually two types of program files: general program files and files on individual projects, clients, or participants.

Focus group: A group of people selected for their relevance to an evaluation that is engaged by a trained facilitator in a series of discussions designed for sharing insights, ideas, and observations on a topic of concern.

History: Events outside the program that affect the responses of those involved in the program.

History and program interaction: The conditions under which the program took place are not representative of future conditions. This is a threat to external validity.

Ideal evaluation design: The conceptual comparison of two or more situations that are identical except that in one case the program is operational. Only one group (the treatment group) receives the program; the other groups (the control groups) are subject to all pertinent influences except for the operation of the program, in exactly the same fashion as the treatment group. Outcomes are measured in exactly the same way for both groups and any differences can be attributed to the program.

Implicit design: A design with no formal control group and where measurement is made after exposure to the program.

Indicator: A specific, observable, and measurable characteristic or change that shows the progress a program is making toward achieving a specified outcome.

Inferential statistical analysis: Statistical analysis using models to confirm relationships among variables of interest or to generalize findings to an overall population.

Informal conversational interview: An interviewing technique that relies on the natural flow of a conversation to generate spontaneous questions, often as part of an ongoing observation of the activities of a program.

Inputs: Resources that go into a program in order to mount the activities successfully.

Instrumentation: The effect of changing measuring instruments from one measurement to another, as when different interviewers are used. This is a threat to internal validity.

Interaction effect: The joint net effect of two (or more) variables affecting the outcome of a quasi-experiment.

Internal validity: The ability to assert that a program has caused measured results (to a certain degree), in the face of plausible potential alternative explanations. The most common threats to internal validity are history, maturation, mortality, selection bias, regression artifacts, diffusion, and imitation of treatment and testing.

Interview guide: A list of issues or questions to be raised in the course of an interview.

Interviewer bias: The influence of the interviewer on the interviewee. This may result from several factors, including the physical and psychological characteristics of the interviewer, which may affect the interviewees and cause differential responses among them.

List sampling: Usually in reference to telephone interviewing, a technique used to select a sample. The interviewer starts with a sampling frame containing telephone numbers, selects a unit from the frame, and conducts an interview over the telephone either with a specific person at the number or with anyone at the number.

Literature search: A data collection method that involves an identification and examination of research reports, published papers, and books.

Logic model: A systematic and visual way to present the perceived relationships among the resources you have to operate the program, the activities you plan to do, and the changes or results you hope to achieve.

Longitudinal data: Data collected over a period of time, sometimes involving a stream of data for particular persons or entities over time.

Macro-economic model: A model of the interactions between the goods, labor, and assets markets of an economy. The model is concerned with the level of outputs and prices based on the interactions between aggregate demand and supply.

Main effects: The separate independent effects of each experimental variable.

Matching: Dividing the population into “blocks” in terms of one or more variables (other than the program) that are expected to have an influence on the impact of the program.

Maturation: Changes in the outcomes that are a consequence of time rather than of the program, such as participant aging. This is a threat to internal validity.

Measurement validity: A measurement is valid to the extent that it represents what it is intended and presumed to represent. Valid measures have no systematic bias.

Measuring devices or instruments: Devices that are used to collect data (such as questionnaires, interview guidelines, and observation record forms).

Micro-economic model: A model of the economic behavior of individual buyers and sellers, in a specific market and set of circumstances.

Monetary policy: Government action that influences the money supply and interest rates. May also take the form of a program.

Mortality: Treatment (or control) group participants dropping out of the program. It can undermine the comparability of the treatment and control groups and is a threat to internal validity.

Multiple lines of evidence: The use of several independent evaluation strategies to address the same evaluation issue, relying on different data sources, on different analytical methods, or on both.

Natural observation: A data collection method that involves on-site visits to locations where a program is operating. It directly assesses the setting of a program, its activities, and individuals who participate in the activities.

Non-probability sampling: When the units of a sample are chosen so that each unit in the population does not have a calculable non-zero probability of being selected in the sample.

Non-response: A situation in which information from sampling units is unavailable.

Non-response bias: Potential skewing because of non-response. The answers from sampling units that do produce information may differ on items of interest from the answers from the sampling units that do not reply.

Non-sampling error: The errors, other than those attributable to sampling, that arise during the course of almost all survey activities (even a complete census), such as respondents' different interpretation of questions, mistakes in processing results, or errors in the sampling frame.

Objective data: Observations that do not involve personal feelings and are based on observable facts. Objective data can be measured quantitatively or qualitatively.

Objectivity: Evidence and conclusions that can be verified by someone other than the original authors.

Order bias: A skewing of results caused by the order in which questions are placed in a survey.

Outcome effectiveness issues: A class of evaluation issues concerned with the achievement of a program's objectives and the other impacts and effects of the program, intended or unintended.

Outcome evaluation: The systematic collection of information to assess the impact of a program, present conclusions about the merit or worth of a program, and make recommendations about future program direction or improvement.

Outcomes: The results of program operations or activities; the effects triggered by the program. (For example, increased knowledge, changed attitudes or beliefs, reduced tobacco use, reduced TB morbidity and mortality.)

Outputs: The direct products of program activities; immediate measures of what the program did.

Plausible hypotheses: Likely alternative explanations or ways of accounting for program results, meaning those involving influences other than the program.

Population: The set of units to which the results of a survey apply.

Primary data: Data collected by an evaluation team specifically for the evaluation study.

Probability sampling: The selection of units from a population based on the principle of randomization. Every unit of the population has a calculable (non-zero) probability of being selected.

Process evaluation: The systematic collection of information to document and assess how a program was implemented and operates.

Program evaluation: The systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future program development.

Program goal: A statement of the overall mission or purpose(s) of the program.

Propriety: The extent to which the evaluation has been conducted in a manner that evidences uncompromising adherence to the highest principles and ideals (including professional ethics, civil law, moral code, and contractual agreements).

Qualitative data: Observations that are categorical rather than numerical, and often involve knowledge, attitudes, perceptions, and intentions.

Quantitative data: Observations that are numerical.

Quasi-experimental design: Study structures that use comparison groups to draw causal inferences but do not use randomization to create the treatment and control groups. The treatment group is usually given. The control group is selected to match the treatment group as closely as possible so that inferences on the incremental impacts of the program can be made.

Random digit dialing: In telephone interviewing, a technique used to select a sample. A computer, using a probability-based dialing system, selects and dials a number for the interviewer.

Randomization: Use of a probability scheme for choosing a sample. This can be done using random number tables, computers, dice, cards, and so forth.

Regression artifacts: Pseudo-changes in program results occurring when persons or treatment units have been selected for the program on the basis of their extreme scores. Regression artifacts are a threat to internal validity.

Reliability: The extent to which a measurement, when repeatedly applied to a given situation consistently produces the same results if the situation does not change between the applications. Reliability can refer to the stability of the measurement over time or to the consistency of the measurement from place to place.

Replicate sampling: A probability sampling technique that involves the selection of a number of independent samples from a population rather than one single sample. Each of the smaller samples is termed a replicate and is independently selected on the basis of the same sample design.

Resources: Assets available and anticipated for operations. They include people, equipment, facilities, and other things used to plan, implement, and evaluate programs.

Sample size: The number of units to be sampled.

Sample size formula: An equation that varies with the type of estimate to be made, the desired precision of the sample and the sampling method, and which is used to determine the required minimum sample size.

Sampling error: The error attributed to sampling and measuring a portion of the population rather than carrying out a census under the same general conditions.

Sampling frame: Complete list of all people or households in the target population.

Sampling method: The method by which the sampling units are selected (such as systematic or stratified sampling).

Sampling unit: The unit used for sampling. The population should be divisible into a finite number of distinct, non-overlapping units, so that each member of the population belongs to only one sampling unit.

Secondary data: Data collected and recorded by another (usually earlier) person or organization, usually for different purposes than the current evaluation.

Selection and program interaction: The uncharacteristic responsiveness of program participants because they are aware of being in the program or being part of a survey. This interaction is a threat to internal and external validity.

Selection bias: When the treatment and control groups involved in the program are initially statistically unequal in terms of one or more of the factors of interest. This is a threat to internal validity.

Setting and program interaction: When the setting of the experimental or pilot project is not typical of the setting envisioned for the full-scale program. This interaction is a threat to external validity.

Stakeholders: People or organizations that are invested in the program or that are interested in the results of the evaluation or what will be done with results of the evaluation.

Standard: A principle commonly agreed to by experts in the conduct and use of an evaluation for the measure of the value or quality of an evaluation (e.g., accuracy, feasibility, propriety, utility).

Standard deviation: The standard deviation of a set of numerical measurements (on an “interval scale”). It indicates how closely individual measurements cluster around the mean.

Standardized format interview: An interviewing technique that uses open-ended and closed-ended interview questions written out before the interview in exactly the way they are asked later.

Statistical analysis: The manipulation of numerical or categorical data to predict phenomena, to draw conclusions about relationships among variables or to generalize results.

Statistical model: A model that is normally based on previous research and permits transformation of a specific impact measure into another specific impact measure, one specific impact measure into a range of other impact measures, or a range of impact measures into a range of other impact measures.

Statistically significant effects: Effects that are observed and are unlikely to result solely from chance variation. These can be assessed through the use of statistical tests.

Stratified sampling: A probability sampling technique that divides a population into relatively homogeneous layers called strata, and selects appropriate samples independently in each of those layers.

Subjective data: Observations that involve personal feelings, attitudes, and perceptions. Subjective data can be measured quantitatively or qualitatively.

Surveys: A data collection method that involves a planned effort to collect needed data from a sample (or a complete census) of the relevant population. The relevant population consists of people or entities affected by the program (or of similar people or entities).

Testing bias: Changes observed in a quasi-experiment that may be the result of excessive familiarity with the measuring instrument. This is a potential threat to internal validity.

Treatment group: In research design, the group of subjects that receives the program. Also referred to as the experimental or program group.

Utility: The extent to which an evaluation produces and disseminates reports that inform relevant audiences and have beneficial impact on their work.

Program Evaluation Resources

Some Web-based Resources

Centers for Disease Control and Prevention: <http://www.cdc.gov/eval/>

Community Tool Box, University of Kansas: <http://ctb.ku.edu/>

Harvard Family Research Project: <http://www.gse.harvard.edu/hfrp/>

Innovation Network: <http://innonet.org>

University of Wisconsin Cooperative Extension:

- Evaluation Resources: <http://www.uwex.edu/ces/pdande/>

- Logic Model Course: <http://www1.uwex.edu/ces/lmcourse>

W.K. Kellogg Foundation: <http://www.wkkf.org/Programming/Overview.aspx?CID=281>

Selected Publications

Connell JP, Kubisch AC, Schorr LB, Weiss, CH. New approaches to evaluating community initiatives. New York, NY: Aspen Institute, 1995.

Fawcett SB, Paine-Andrews A, Francisco VT, Schulz J, Ritchter KP, et al. Evaluating community initiatives for health and development. In: Rootman I, Goodstadt M, Hyndman B, et al., eds. Evaluating Health Promotion Approaches. Copenhagen, Denmark: World Health Organization (Euro), 1999 (In press).

Fawcett SB, Sterling TD, Paine Andrews A, Harris KJ, Francisco VT, et al. Evaluating community efforts to prevent cardiovascular diseases. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, 1995.

Fetterman DM, Kaftarian SJ, Wandersman A. Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage Publications, 1996,

Patton MQ. Utilization-focused evaluation. Thousand Oaks, CA: Sage Publications, 1997.

Rossi PH, Freeman HE, Lipsey MW. Evaluation: A systematic approach. Newbury Park, CA: Sage Publications, 1999.

Shadish WR, Cook TD, Leviton LC. Foundations of program evaluation. Newbury Park, CA: Sage Publications, 1991.

Taylor-Powell E, Steele S, Douglas M. Planning a program evaluation. Madison, WI: University of Wisconsin Cooperative Extension, 1996 (see Web-based entry on page 66).

University of Toronto, Health Communication Unit at the Center for Health Promotion. Evaluating health promotion programs (see Web-based entry on page 66).



www.cdc.gov/getsmart