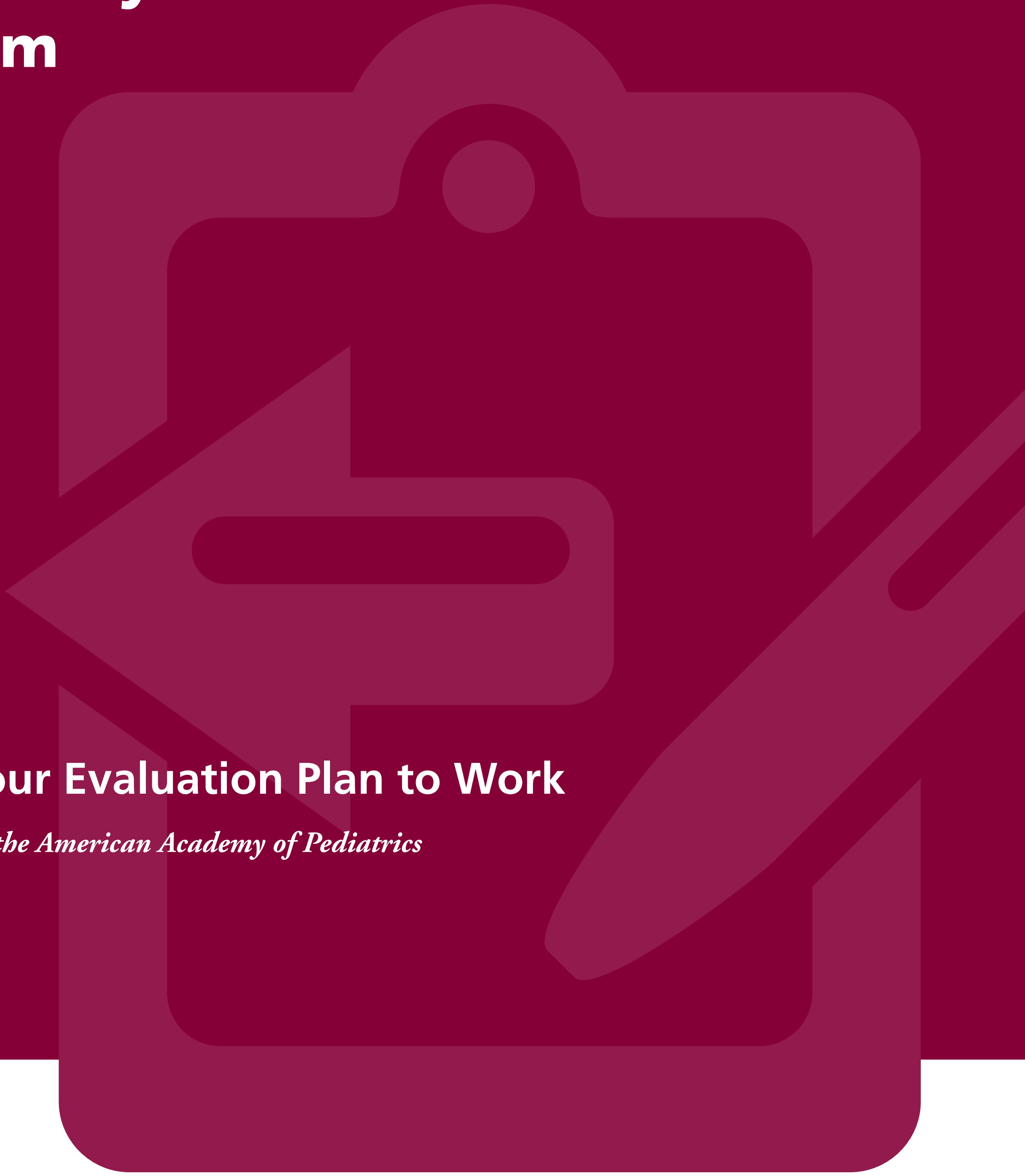
Evaluating
Your Community-Based Program



Part II

Putting Your Evaluation Plan to Work

A publication of the American Academy of Pediatrics

HEALTHY TOMORROWS
Partnership for Children

A Cooperative Agreement Program of the Federal Maternal and
Child Health Bureau and the American Academy of Pediatrics

American Academy of Pediatrics
DEDICATED TO THE HEALTH OF ALL CHILDREN™

Acknowledgments

The American Academy of Pediatrics (AAP) would like to thank the Maternal and Child Health Bureau at the Health Resources and Services Administration in the Department of Health and Human Services for its support and leadership in implementing the Healthy Tomorrows Partnership for Children Program. We would like to acknowledge the District Community Access to Child Health Facilitators for their ongoing oversight of the program and for reviewing the manual. We would also like to thank the additional reviewers: Karla Palmer, Lynn Olson, Burke Eilers, Jim Tehan, Alina Bueno Nadsady, and Richard Wilson. Lastly, we would like to give special thanks to the authors of this guide: Holly Ruch-Ross, Nita Patel, and Nicole Miller.

This publication was supported by the American Academy of Pediatrics (AAP) Healthy Tomorrows Partnership for Children Program/Cooperative Agreement # U50MC07618 from the Maternal and Child Health Bureau (MCHB). Its contents do not necessarily represent the official views of the MCHB or AAP.

www.aap.org/EvalResources

4-79/1208

Table of Contents

Introduction

Community-based programs and their community partners are increasingly interested in evaluating the work that they do in their communities. The interest may be externally driven, such as funders' requirements or the need to prove "value" to community leaders, or it may be internally motivated, such as the need to obtain new funding to expand or sustain services or simply the desire to explore program effectiveness and improve quality. There is growing evidence that program sustainability is supported by carefully planned and well-implemented ***program evaluation***. However, few community-based programs begin with the resources or expertise to conduct a good program evaluation.

This publication is the second of a two-part guide to program evaluation developed by the American Academy of Pediatrics (AAP) for Healthy Tomorrows Partnership for Children Program grantees and applicants. It is also intended to be useful to pediatricians and others implementing community-based programs. The purpose of the guide is to provide quick overviews of major issues in program evaluation and to point you toward the broad array of resources for high-quality program evaluation that are available. After reading *Evaluating Your Community-Based Program—Part II: Putting Your Evaluation Plan to Work*, you will be able to:

1. Identify ways to measure progress on your goals and objectives.
2. Select tools and strategies for collecting information you need to evaluate your program.
3. Analyze and present your information in a meaningful and useful way.
4. Know where to go for additional information on these topics.

Part I of this guide, *Designing Your Evaluation*, focused on understanding and planning a good evaluation. Part II emphasizes effective documentation to evaluate your program. It is also intended to help you decide how to measure progress on your objectives and collect, analyze, and present the resulting data meaningfully and efficiently.

The guide is structured in a workbook format, so there is space to apply each concept to your project as you go along. Each section also includes a case study example to demonstrate how evaluation ideas within a single program will develop over time.

We have included a glossary and appendix of additional resources at the end of this installment. Terms that appear in ***bold italics*** throughout this guide are defined in the glossary. We've also included **Jargon Alerts** in relevant sections to help you understand unfamiliar terms.

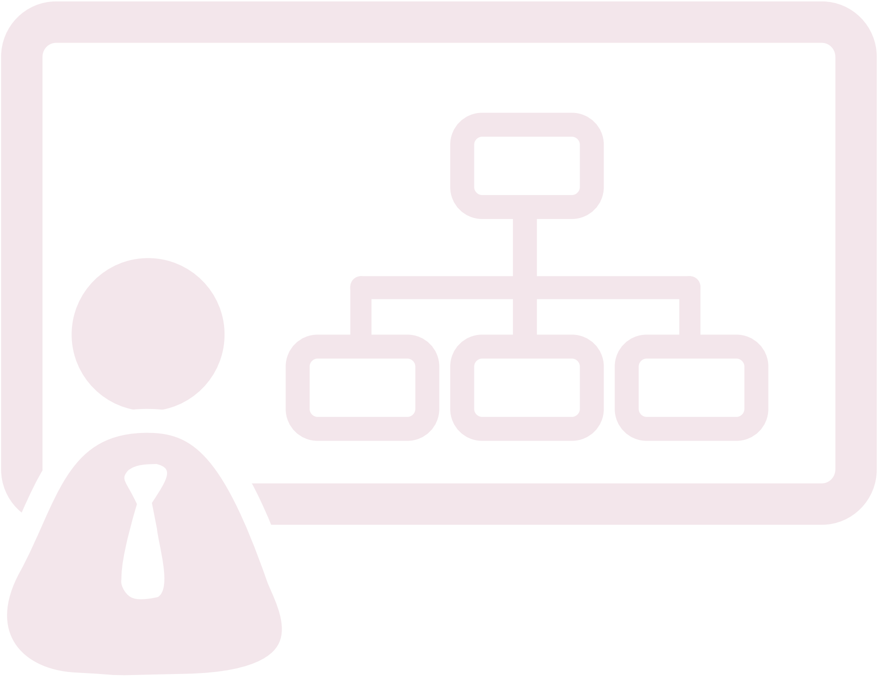
A Quick Review of Part I

Evaluating Your Community-Based Program: Designing Your Evaluation

Part I of this guide is titled *Evaluating Your Community-Based Program: Designing Your Evaluation*. In this first volume, readers learned to:

1. Understand the roles evaluation plays in program design and improvement.
2. Understand the importance of stakeholder input and involvement in evaluation design.
3. Define the outcome(s) a program plans to accomplish.
4. Complete a logic model for a program.
5. Know where to go for additional information on these topics.

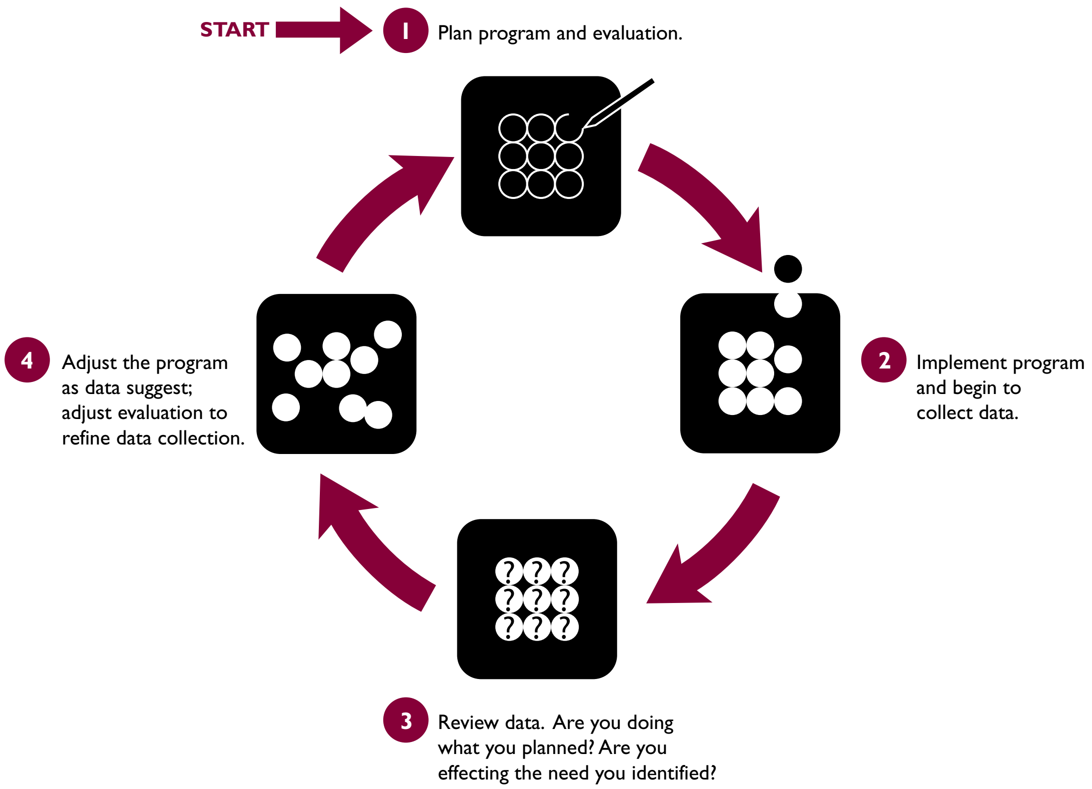
We recommend reviewing Part I of *Evaluating Your Community-Based Program* guide as a foundation for the information in Part II. It is available on the Web site of the American Academy of Pediatrics, www.aap.org/EvalResources or by either calling the AAP Division of Community-based Initiatives at 847/434-7085 or e-mailing healthyt@aap.org.



What Is Evaluation?

For purposes of this guide, *evaluation* is defined as an *ongoing* process of systematic assessment of your program to provide information needed for making decisions. Evaluation is distinct from other forms of research in that it is action-oriented, seeking to provide information that is useful for decision making relevant to program development and improvement. We see evaluation as a cycle:

Figure I: Evaluation Cycle



You will often see a distinction drawn between process evaluation and outcome evaluation. ***Process evaluation*** answers the question, "Are we completing the ***activities*** or providing the services that we believe lead to the changes we desire?" In other words, *is your program proceeding as planned*? Process evaluation examines the implementation of your program. ***Outcome evaluation*** seeks to answer the question, "Is my project making a difference?" ***Outcomes*** try to describe the ***impact*** of a program on a community beyond the point of service. In our view, both aspects of evaluation are integral to good program management.

Why Evaluate?

Although programs vary in their need for information and the decisions that program managers must make, most programs will have the following evaluation needs:

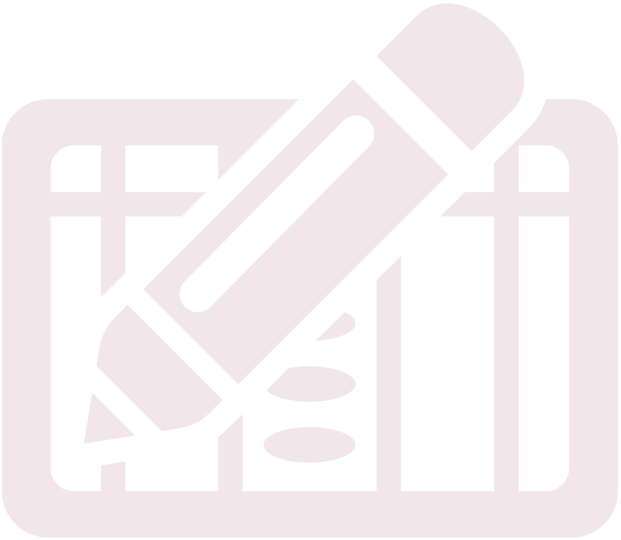
1. **Check Your Process:** Confirm that you are doing what you said you would do.
2. **Determine Your Impact:** Check that you are having the desired effect in the target population.
3. **Build Your Base of Support:** Generate information and evidence to share with funders and other stakeholders.
4. **Justify Replication:** Find evidence to support the expansion or replication of this program.

Almost everyone who has a vision for your project and is interested in finding out whether it "works" should be involved in planning your evaluation.

The Evaluation Plan: Goals, Objectives, and Outcomes

A good evaluation plan begins with clear ***goals*** that describe what you want to achieve. Your goal statement articulates what your program would like to accomplish in relation to a specific problem and target population. Once you are clear about the goal or goals of your program, you can define your ***objectives***, or the steps you will take in your efforts to achieve your goal. Good objectives are SMART: **S**pecific, **M**easurable, **A**chievable, **R**ealistic, and **T**ime specific.

Clear goals are essential for both program implementation and evaluation, but tracking your progress requires identification of outcomes. Outcomes are measurable changes that occur beyond the point of service or intervention that reflect the logical and desired result of your services.



The Logic Model

The ***logic model*** is a tool that has been adapted from a business model developed to explain the logical relationship from strategy to return on investment. It is widely used in social service fields and by some government agencies to facilitate program planning, implementation, and evaluation. Your logic model provides a snapshot of your program and serves as a single-page summary of your program that is easily shared with staff, boards, and funders. **The development of a logic model is in itself often a valuable consensus-building process.**

	Target Population	Inputs	Activities	Outputs	Outcomes
IDEAS	The characteristics of people or communities you work with and the needs they present	The resources required for this program to operate	Strategies you use or services you provide to try to achieve your goal	Basic data on program participation	Desired changes in the target population as a result of program activities
EXAMPLES	*Age, gender, socioeconomic status, ethnicity, language, geographic location, low health care use, high cancer incidence, lack of mental health information, etc.*	*Money, staff, volunteers, facilities, etc.*	*Provide training, counseling, education, screenings, referrals, develop materials, etc.*	*Number of participants attending training, number of counseling sessions, etc.*	*Changes in knowledge, attitude, behavior, health status, health care use, incidence, prevalence, etc.*

What's Next?

This second part of *Evaluating Your Community-Based Program* is intended to take you from your logic model and evaluation plan through the process of identifying measures, selecting strategies for collecting and analyzing information, and presenting your results.

In *Designing Your Evaluation*, we met Sarah, the director of the Prevention First Program. Sarah and her program went through the same process you need to go through to specify your program and develop your evaluation plan.[1]

CASE STUDY

Sarah is the program director of the Prevention First Program, located at a large multiagency collaborative in an immigrant community. The community experiences high mobility, is very low income, speaks limited or no English, and has low rates of utilization of preventive health care services. Residents have come to this country from a variety of cultures. Sarah's program intends to bring together the variety of resources and expertise present in the collaborative to try to facilitate the use of preventive health care by this community and to increase public awareness of the many free, nonemergency health and dental services available in the community.

In Part I of this guide, we followed Sarah and her staff as they developed the goals and objectives of the Prevention First Program.

Prevention First Program Goals:

1. *Immigrant families will understand the importance of prevention.*
2. *Immigrant families will use preventive health services.*

Prevention First Objectives:

1. *Within the first 6 months of the project, we will conduct a focus group with immigrant parents to explore possible barriers to the use of prevention services.*
2. *By the end of year 1, we will have made presentations to staff of at least four agencies serving immigrant families to promote preventive health services and encourage referrals.*
3. *By the end of year 1, participating immigrant families will schedule and complete an increased number of well-child visits over baseline.*

*Based on these goals and objectives, Sarah and her staff developed a logic model for the Prevention First Program. The logic model summarizes the program by specifying the target population and identified needs, the **inputs** and activities of project staff, and both program outputs and anticipated outcomes for the target population. This logic model serves as the foundation for evaluation planning for the Prevention First Program.*

---

[1] This case study represents a hybrid of common experiences among many projects, but it is fictional. Any similarity to an actual project is purely coincidental.

CASE STUDY

Prevention First Logic Model

Target Population and Needs	Inputs	Activities	Outputs	Outcomes
• Low-income, limited English-speaking immigrant community, with Spanish, Polish, Portuguese, and French as the primary languages • Low use of health care coverage • Low use of preventive health services • Living primarily within the same 4 square mile area known as "First Stop" • Mostly employed in temporary and/or part-time labor positions • Mostly from cultures without preventive medicine or health care practices	• Coalition members, director, and two interns • Funding (foundation, local, and state) • Computers and software • Prevention education curriculum and volunteer health educators • Prevention media • Verbal and written translation	• Health care use intake, collected verbally and language appropriate: brief format in nontraditional settings or full intake in traditional settings • Complete registration for health care coverage in community settings for new and lapsed families • Prevention education sessions held in community settings • Preventive health services offered regularly in nontraditional community locations • Conduct focus groups • Regular tracking of health care coverage and preventive service use	• Number of new families signed up for health care coverage monthly • Number of lapsed coverage renewed monthly • Number who attended prevention education monthly • Number of preventive health care services contacts in nontraditional settings • Number of preventive health services contacts in traditional (clinic) settings • Number of focus groups conducted	• Immigrant families will understand the importance of preventive health care services. • Participating immigrant families will schedule and complete an increased number of well-child visits. • Immunization rates will increase among children in the target population. • The number of workdays or school days missed due to illness will decrease.

We will continue to follow Sarah in *Putting Your Evaluation Plan to Work*, as she and her staff make decisions about measurement, data collection, analysis, and presentation of information about her program.

After the Logic Model: Gathering Information

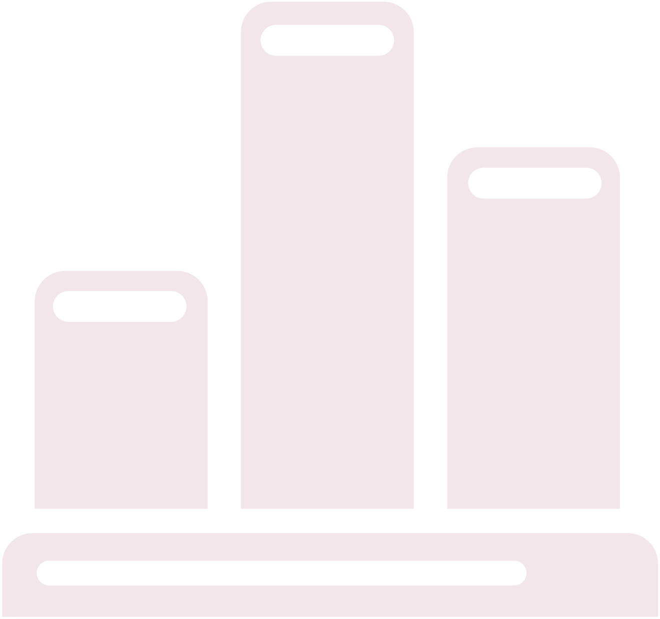
Deciding What Information You Need to Gather

Virtually all programs need to gather information to be able to serve their target populations and their communities. If your logic model reflects your program well, it already lists much of the information you need to collect. You will need information to serve two broad purposes: documenting process and assessing outcomes.

Documenting process. Keeping track of what comprises the program is fundamental to program evaluation. Process evaluation allows you to describe your program, which makes outcome evaluation meaningful and replication of the program possible. Usually, program staff know that they need to keep track of the services they are delivering. This is the beginning of documenting program process. Process documentation will help you answer questions such as:

- Is program implementation going as planned?
- How many people are receiving services?
- Who is receiving services?
- What services are people receiving?
- How many people are we referring to other providers?

Process documentation helps you describe what you are doing and with whom you are doing it. It is also essential for fine-tuning your program. The most useful process documentation is collected in a consistent manner using a well-planned system of data collection. If your logic model reflects your program well, much of the information you will want to collect for your process evaluation will be listed in the third and fourth columns as activities and ***outputs***.



CASE STUDY

Sarah and the Prevention First Program staff knew that process evaluation would be critical to understanding the results of their interventions. They used their logic model to list the activities and outputs they needed to document and their plan for documenting each. As they developed the record-keeping system and forms they would use, the staff ensured that each of these items was included.

Activities	Outputs	Documentation Plan
Health care use intake, collected verbally and language appropriate: brief format in nontraditional settings or full intake in traditional settings	Number of completed intake forms	Count number of forms completed
Complete registration for health care coverage in community settings for new and lapsed families	Number of new families signed up for health care coverage monthly; number of lapsed coverage renewed monthly	Case notes; count coverage applications
Prevention education sessions held in community settings	Number who attended prevention education monthly	Group participation logs (sign in)
Preventive health services offered regularly in nontraditional community locations	Number of preventive health care services contacts in nontraditional settings	Case notes; contact forms
Focus groups	Number of focus groups conducted	Count focus groups
Regular tracking of health care coverage and preventive service use	Number with access to services and use of services	Tracking system—family health care coverage; use of preventive services

WORKSPACE

Now it's your turn to identify the process information that you must collect in order to document your program. You can fill in your activities and outputs from your logic model, and then think about what you have or need to collect for each.

Activities	Outputs	Documentation Plan

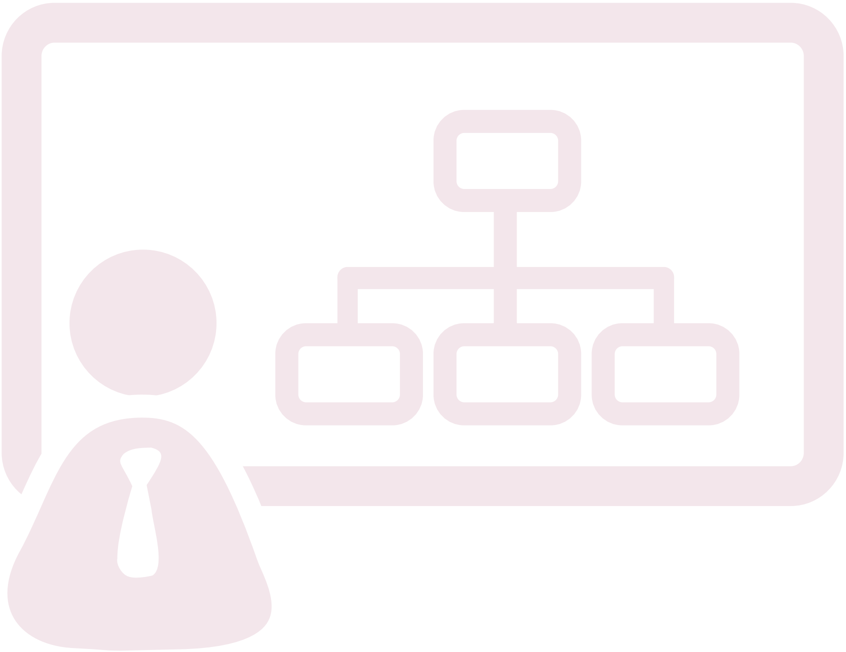
SIDEBAR

So You Thought You Were Done with Needs Assessment?

Documenting your program may involve additional or even ongoing community assessment. When you wrote a grant proposal to fund your project, you presented information about your community to establish the need for the services you were proposing. But this is not the end of our need for community assessment. As your program evolves, you will find yourself asking questions such as these:

- How do I let my target population know about the services we offer?
- How can I develop trust with the people in my target population?
- What strengths (assets) in the families in this community will facilitate use of our services?
- Are there language, transportation, or other barriers that, unaddressed, will limit our target population's use of our services?

These and other questions related to **what services are needed** and **how to get needed services to the right people** require ongoing community assessment. This may take the form of finding existing information compiled by someone else (such as your health department or local school district) and will also likely include informal information gathering (such as reading the local newspaper or talking to community leaders). It may also require structured data collection by your program (such as household surveys, ***focus groups,*** and observations at community gatherings). Most of this kind of information will bear directly on the first column of your logic model, where you identified your target population and its needs.



Assessing outcomes. You are trying to effect a change in your community; this is the reason you designed your program in the first place. In the fifth column of your logic model, you have listed your outcomes, or the desired changes in the target population. To see whether your program is succeeding, you will need to know whether these outcomes are being achieved. This means that you will need a way to measure the changes that you hope are occurring. This is the focus of your outcome evaluation.

For some common problems there are standards in the field for what is considered a positive outcome. For example, in a program to reduce the ***incidence*** of childhood lead poisoning, a reduction in blood lead levels in your program participants to within acceptable standards would be a certain indicator of success. If your program is working to address this type of issue, then your measure of success has already been defined for you. Your measurement challenges will involve ensuring that measures are timely and appropriately administered, but you will not have to struggle with the question of what to measure.

Unfortunately, outcomes are not always easy to measure directly. Desired outcomes are often longer term than the project itself or involve concepts that are not directly observable. It may be simply unrealistic, within the context of a given program and the resources available, to measure the desired outcomes. In these circumstances, we often have to settle for measurement of an indicator of progress toward the achievement of our desired outcome. ***Indicators*** are approximations of desired outcomes that can be measured when the outcome cannot.

Like good objectives, indicators need to be SMART (specific, measurable, achievable, realistic, and time specific). An indicator may represent an intermediate step toward the achievement of an outcome, such as school attendance as an indicator of progress toward high school graduation. Often, an indicator is an effort to capture a construct that is not directly observable, such as self-esteem or anxiety.

JARGON ALERT

!

Indicator

Indicator is a measurable intermediate step or other approximation of an outcome. An indicator is used when the outcome itself is difficult to measure directly and/or difficult to measure during the time frame of the project.

CASE STUDY

Sarah and the Prevention First Program staff included four outcomes in their logic model. Sarah examined the literature and talked to both colleagues and people in her community. Using the information available in their field, the Prevention First Program staff identified an indicator for each of the outcomes.

Prevention First Program Outcomes and Indicators

Outcome	Indicator
Immigrant families will understand the importance of preventive health care services.	Family rating of the importance of preventive health care after 6 months of program participation
Participating immigrant families will schedule and complete an increased number of well-child visits.	Number of well-child visits among children from participating families in the first and second years of the program
Immunization rates will increase among children in the target population.	Change in immunization rate 2 years after program is implemented
The number of workdays or school days missed due to illness will decrease.	Participant report of missed days after 1 and after 2 years of program participation

WORKSPACE

Now it’s your turn to think about indicators for your program outcomes. Can you identify a measurable indicator for each of the outcomes your program seeks to accomplish?

Outcomes and Indicators

Outcome	Indicator

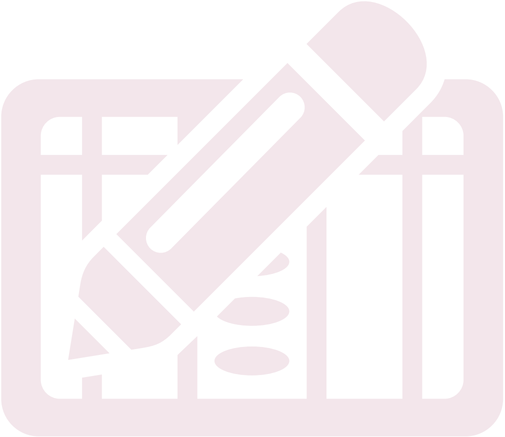
SIDEBAR

Are you having trouble keeping your outputs, outcomes, and indicators straight?

Output	**Outcome**	**Indicator**
Direct products of program activities	Changes in the target population that result from the program	Intermediate step or approximation of an outcome
Reflect program implementation	Reflect program impacts	Reflect steps toward program impacts or proxy measures
Example: Parents attend an oral hygiene training session.	Example: Young children are free of dental caries.	Example: Parents have knowledge of oral hygiene practices.

Comparison information. Another area in which programs are frequently challenged is in identifying good ***comparison information.*** If you are hoping to document change in an indicator, you may need to collect ***baseline*** information on participants upon their entry into the program. This baseline information provides something against which you can compare participant outcomes after they have participated in the program. Although not generally regarded as conclusive proof, participant change in the desired direction provides some evidence for your program's effectiveness.

The "gold standard" for comparison is the ***randomized clinical trial,*** which controls for many of the factors that undermine your ability to make the case for the effectiveness of your intervention.



JARGON ALERT

Comparison information

Comparison information is information against which program information can be compared, usually to assess whether program participants are in some way different from other populations or to assess their status after program participation. Participants might be compared based on their characteristics or on some indicator of progress or outcome.

JARGON ALERT

Baseline

Baseline refers to information about the target population or community collected before the intervention begins. Baseline information on indicators and outcomes is useful in examining whether there has been a change in the target population.

JARGON ALERT

Randomized clinical

A ***randomized clinical*** trial is a research study that utilizes a pool of subjects from a population of interest who are randomly assigned to treatment and control (no treatment) conditions.

But randomly assigning participants to treatment (your program) and no treatment conditions is challenging in most service delivery settings, and following a control group is simply beyond the resources of most program budgets. Although none meet the standard of the randomized clinical trials, programs may have other options for obtaining some comparison information:

- Pretesting and posttesting, which usually test for changes among program participants in knowledge, attitude, or self-reported behavior (Each program participant serves as his or her own control.)
- A ***convenience sample,*** which uses another group in the community on whom data are or can be collected to compare with program participants
- National, state, or local data
- A benchmark, or absolute standard, against which program results can be compared

The great benefit of having comparison information is that it controls for some of the factors that may affect your results. Of particular salience in the case of a community-based program are:

- **History,** or things that happen in your community outside of your project. For example, if a new state law makes it more difficult for families to prove eligibility for a service you provide, that will almost certainly affect your program and its participants.
- **Passage of time,** or the natural maturation process that occurs over time. Natural maturation is almost always an issue for programs that provide services to children. In a perhaps slightly silly example, rarely will anyone be surprised or impressed if your program is able to prove that children's height increases after a year of program participation!
- **Selection bias,** the bias introduced based on who is missed in your service delivery or data collection. A brief telephone follow-up survey of program participants, for example, will miss anyone who does not have a telephone or who has recently changed his or her phone number. It is not always immediately obvious when selection is occurring, but it is often the explanation for findings that are odd or difficult to explain.

Source: Gathering Information on Your Outcomes

When you start to think about the information you need to collect, it's a good idea to think first about what you already have.

Information you already have. If your program already exists, or if you are a part of a larger organization, there is a good chance that, like the Prevention First Program staff, you have program records that already collect some of the information you will need. Think about any charting, summaries, or reporting that you already do that document information such as number of contacts with the project. Maybe you already have much of your process data collection set up! You may need to tweak some of your forms and protocols to meet your program's needs, and this could be a good opportunity to simplify or streamline your data systems. In addition, if your program is an expansion or enhancement of existing services, your organization may already have collected information about some of the people who will be participants in the program. Can you access existing information such as family background and medical history rather than collecting it again? You may also already have information about your community, and you may even have data that can be used as a baseline or comparison on an outcome of interest.

Existing data in your community. Many organizations collect data, particularly on a community level. For example, your local, county, or state public health department or the local organization you collaborate with may have community data that provide baseline or comparison information about your outcomes. For the Prevention First Program, Sarah and her staff realized that the community-level immunization data collected by their county would be useful in assessing their program outcomes. Sharing individual-level information is more problematic, because organizations carry a heavy responsibility to protect the privacy of the people with whom they work. The US government has established legal requirements related to privacy and protection of human subjects (see Appendix B for information on the Health Insurance Portability and Accountability Act [HIPAA] and Institutional Review Boards [IRB]).

Information you will need to gather. You will almost certainly discover some need for additional data collection. In addition to knowing what information you need, you will also need to make some decisions about who has the information, from whom you will collect it, and the time frame in which it will be collected. Your information gathering system needs to target the best source of information needed for your program, and it has to be organized in a way that allows that source to easily share what he or she knows. Next, think about how you will best be able to get the information. Many strategies for collecting information can be used, including individual and group approaches; face-to-face, telephone, mail, e-mail, and Internet approaches. Each strategy has its advantages and disadvantages, and some fit better within a given program setting than others. Moreover, if program staff will conduct much of the data collection, some strategies will be more appropriate than others. For example, an interview that asks directly about the effectiveness of services cannot be credibly conducted by the individual who has delivered those services. A staff person who has worked with and grown to know a family cannot realistically be expected to conduct objective observations of interactions within that family.

If it is feasible, using more than one strategy can strengthen the overall findings, as one method may help compensate for the limitations of another. Some of the more familiar methods are described in the following table.

Data Collection Methods

	Description	Uses	Limitations
Interview	The interviewer asks a series of questions of the subject and records responses. Interviews vary in the level of structure. A single interview may include assessment of multiple domains. Data may be quantitative, qualitative, or a mix of the two.	Interviews are good for obtaining individual-level information not easily collected on a written questionnaire. They are appropriate for individuals unable to complete a questionnaire. Interviewing also allows for in-depth probing and follow-up questions.	Interviewing is labor-intensive. An interview requires that the subject be comfortable with and able to trust the interviewer; he or she may feel some loss of privacy.
Questionnaire or survey	It consists of a series of written questions to which the subject responds. Questions may include open-ended, short answer, and forced choice questions. Like an interview, a questionnaire may include many different scales within one instrument. If resources permit, a questionnaire may also be administered electronically. Data may be quantitative, qualitative, or a mix of the two.	Questionnaires allow for anonymous response. They are generally cost-effective.	A questionnaire can be burdensome for certain categories of respondents and may yield problematic data if respondents do not understand the questions in the same way as the question designer does. A questionnaire is not useful when categories of response cannot be anticipated, and it is not as well suited to "why" or "how" kinds of questions.
Knowledge assessment	A knowledge assessment is a test of knowledge, commonly seen as a pretest before an intervention and a posttest afterward. Data are usually quantitative.	Pretests and posttests are commonly used as quick, inexpensive indicators of program progress.	Questions must be well chosen to sample the content of the intervention and must be appropriate for the culture and literacy level of the target population. For most programs, knowledge gains are necessary but not sufficient outcomes.
Biometric test	A biometric test is a physiologic measurement, such as body mass index (BMI), blood lead level, or bone density. Tools such as a pedometer or accelerometer also measure a physiologic function. Data are usually quantitative.	Usually, biometric indicators are highly reliable and well established for certain conditions.	Biometric indicators are often relatively expensive to collect.
Observation	A trained individual (or individuals) observes an environment or process using a specific protocol for recording or rating incidents of interest. Data may be quantitative, qualitative, or a mix of the two.	Observation can be useful for assessing change in an environment or for understanding a process.	Observation is dependent on the expertise, objectivity, and consistency of the observer(s).
Chart review	Information is systematically extracted from existing patient or client records. Data may be quantitative, qualitative, or a mix of the two.	Information has already been gathered, so data can be extracted without concern for subject scheduling, etc. Any sampling method can be used, including a random sample.	Chart review is somewhat labor-intensive, and missing information can be a particular problem.
Focus group	A focus group is a professionally facilitated, focused discussion among a group of people to help understand a topic of interest. Usually, individual focus group participants are selected based on specific characteristics relevant to the topic. Data are usually qualitative.	Focus groups can be excellent for identifying concerns, needs, and barriers, as well as for providing richer understanding of what goes on in the target population.	Focus groups do not necessarily represent the dominant experience of the population. Meeting the standards for a true focus group can be somewhat challenging for programs. Language barriers are particularly salient for focus groups; the facilitator must be fluent in the participants' language, and participants who speak different languages cannot be combined.

SIDEBAR

A Note about Online Surveys

The use of the Internet for data collection is a relatively recent phenomenon. In an online survey, members of the target group receive an e-mail message asking them to go to a Web address (usually, by clicking a link) and complete an electronic questionnaire. There are some great advantages to this method; however, the disadvantages are also substantial.

Advantages of Online Surveys	Disadvantages of Online Surveys
• Fast	• Limitations on questions and formats
• Inexpensive	• Requirement of reliable Internet access for target group
• Automated process, including compilation of data	• Potentially poor response rate

CASE STUDY

Sarah and the Prevention First Program staff identified an indicator for each of the four outcomes in their logic model. Now they need to determine a potential source of information for each of the indicators selected.

Prevention First Program Sources

Outcome	Indicator	Source
Immigrant families will understand the importance of preventive health care services.	Family rating of the importance of preventive health care after 6 months of program participation	Participating parents—possible survey
Participating immigrant families will schedule and complete an increased number of well-child visits.	Number of well-child visits among children from participating families in the first and second years of the program	Program records—family case records, visit logs
Immunization rates will increase among children in the target population.	Change in immunization rate 2 years after program is implemented	Community-level immunization data from the county
The number of workdays or school days missed due to illness will decrease.	Participant report of missed days after 1 and after 2 years of program participation	Participating parents—possible survey, school records

WORKSPACE

Now it's your turn to think about sources for your program indicators. Can you identify a source for each of the indicators of your program?

Outcomes and Indicators

Outcome	**Indicator**	**Source**

Remember that timing is everything. Other than basic demographics, most of the information you will collect will have some time sensitivity. Be sure to pay attention to this in scheduling data collection. In addition, information that is being collected from program participants should ideally be tied to the sequence of program activities.

Selecting the Right Tools

You have identified the information you will need to gather and identified from whom you will gather it. The next step is to determine which data collection tools are appropriate. This is a process that may take some time, so the project timeline needs to take into consideration the possibility that tools will need to be created or adapted. In the selection process, a tool has to provide the proper information needed, the source of the information, and the person (usually, a staff member) who will collect it.

You may be able to use a tool that someone else has already created, or you may decide that you need to develop a tool that is specific to the needs of your program. For much of your process data collection, for example, you may use tools provided by your organization or developed for related programs to document program activities and outputs. You may need to develop or adapt some tools or forms to meet the specific needs of your program. In any case, there are some important considerations in selecting the tool or tools that you will use to collect information. A good tool should be:

JARGON ALERT

!

Tool

A ***tool,*** in evaluation and research, is a specific mechanism used to collect information; it is also sometimes known as an instrument or a measure. Examples of types of tools include questionnaires, clinical checklists, computer programs, and observational rating scales.

Simple. Record-keeping that feels burdensome to either staff or program participants is less likely to be kept current or completed at all. Collect only what is needed, and minimize the requirement for narrative.

Realistic. Think twice about including information on standard forms that is difficult to obtain or very sensitive for program participants. If information is spotty, it will be hard to interpret and use later.

Used consistently. Everyone who is going to use a tool or form needs to understand what information is being gathered, when it is gathered, and what its intended use is. Ideally, your entire program staff will be involved in designing a tool; at a minimum, they will need to be trained together to ensure that everyone has the same understanding of the questions being asked and the possible responses.

In a useful form. Questions that can be answered by checking a box or filling in a number are simpler both to answer and to compile later. Short answer questions are much easier to handle than a long narrative.

A measure of the right construct. Even a very good instrument isn't useful if it measures the wrong thing. This is a particular issue in selecting measures of outcome. Be firm with yourself (and those who advise you!) about your outcomes and appropriate indicators.

Appropriate for the target population. Factors related to age, culture, language, and other issues will affect how well you are able to collect data. If you would like program participants to complete a questionnaire, for example, the language spoken and the level of literacy in your target population must be considered. Tools developed for adolescents may not be well suited to assessing the same construct in younger children. Some groups may not be able to complete a self-administered questionnaire in a reliable or consistent manner.

Easy to administer. Much of your data collection will likely occur during program activities. Information will need to be gathered in the places in which the program takes place, such as a home or an office, or in a group setting. Time constraints and factors such as the presence of young children need to be considered.

Using an Existing Tool

There are many advantages to finding an existing tool that addresses these issues and meets your program's needs. It gives you some security about what you are measuring and lends credibility to your findings. It may provide the opportunity to compare your results to existing data on similar or different populations and programs. A well-established tool is particularly valuable when you want to assess a construct that is not directly observable, such as self-esteem. One caution, however: even the most elegant tool is not useful if it does not measure the outcome or indicator of interest.

Finding Existing Tools

Program managers often find themselves overwhelmed by the task of finding relevant tools that already exist. This can, indeed, be a daunting undertaking, particularly as measurement is its own separate area of inquiry, with its own experts and distinctive literature and jargon. However, it is not necessary to be a measurement expert to find a useful tool (although some detective skills might be helpful). Consider tools being used by programs that are similar to yours or that have similar goals and objectives. You may learn about these tools in the professional literature or through contact with other programs and professionals. These tools will be doubly valuable, because they have already been tested in a service delivery setting.

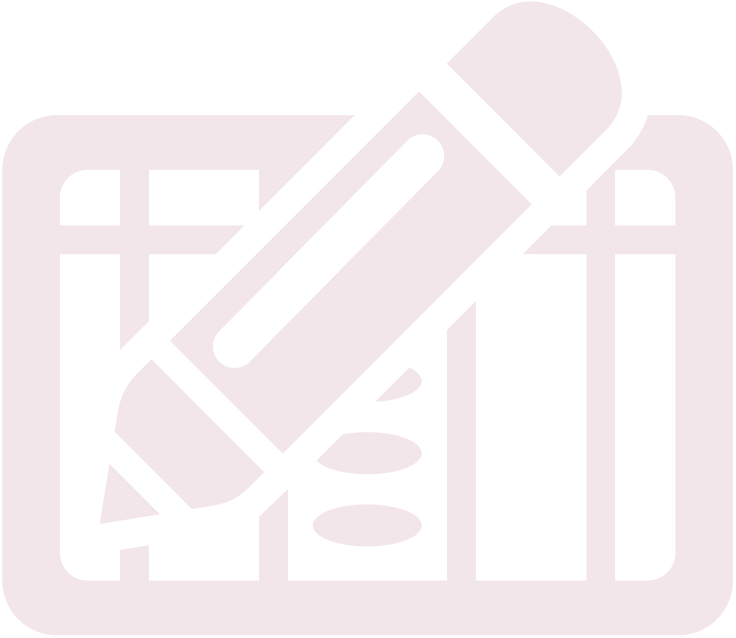
Published measures of specific disciplines. Before the Internet, these usually thick books were worth their weight in gold, offering basic information and reviews of collections of tools. The books are still being published because, even with the availability of information on the Internet, a book with the right focus sitting on your shelf can save you many hours of searching. See, for example, *Measuring Health*, by Ian McDowell and Claire Newell[2]; *Measures for Clinical Practice, Volumes 1 & 2*, by Joel Fischer and Kevin Corcoran[3]; or one of the resources published by the American Psychological Association. In some cases, a compendium of measures comes from a single project or set of related projects.

[2] McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 3rd ed. Oxford: Oxford University Press, 2006.
[3] Fischer J, Corcoran K. *Measures for Clinical Practice: A Sourcebook*. 4th ed. New York, NY: The Free Press, 2006.

Tools on the Internet. Many resources for identifying tools have moved to the Internet. To some extent, success in finding tools on the Internet is contingent on good search skills, but measures are more likely to turn up on certain kinds of Web sites than on others. One very fruitful area to pursue is sites operated by organizations or institutions that develop and utilize the tools. Often, these are university-affiliated groups who use the tools as a part of their professional work. One example is the Community Toolbox Web site operated by the University of Kansas; it includes measures of community change among its tools to promote community health and development. Sites operated by organizations that support research and evaluation (often foundations or federal agencies) can be particularly helpful. Another potential source is sites operated by interest or advocacy groups, such as the American Diabetes Association, which sometimes include measures or links to relevant research sites. Often, once you find one relevant site or tool, you are on your way to finding others.

Appendix A to this guide includes full references for the resources mentioned and additional examples of resources for finding tools.

If you are using an existing tool, you will also want to be sure that it is readily available, affordable, and, preferably, supported by the author. If the author supports the tool, he or she may be able to offer additional information and advice in using it in your target population. Some tools are proprietary, meaning that you must have the author's permission to use them. This can be an advantage in that it means the author continues to support the tool and wants to know how it is being used and how it works in diverse populations and settings. However, it may also constrain your ability to use a part of an instrument or to adapt it to your population. You may be unable to select a particular tool simply because it is too costly for your program budget.



SIDEBAR

Assessing the Quality of Published Measurement Tools

Tools that are reported in the professional literature have usually been tested on several dimensions used to assess the quality of measures. When you have identified potential tools, information about their validity, reliability, and standardization may be available to help you assess their quality.

Validity. ***Validity*** is a representation that a tool really measures what it is intended to measure. Validity is assessed in many ways, including something as simple as expert opinion. Another very common way to assess validity is to compare the results of a new tool with those obtained using a standard tool to see how closely the results align (for example, a new paper-and-pencil test compared with a clinical assessment of a psychiatric condition). Validity is also frequently examined by testing whether a new measure behaves as we would expect in relation to other things that are known about the group of people on whom it is being tested.

Reliability. The ***reliability*** of a tool refers to both its consistency across multiple administrations and to the internal consistency of the items that comprise the measure. A tool that is reliable is one that includes items that are related to one another and that give a consistent estimate of the construct being measured. For example, in measuring depression, if three questions all point to depression, then the measures report consistent findings.

Standardization. ***Standardization*** means that a tool has been tested in one or more populations and the results are consistent across the groups. When a tool is standardized, you can get information that helps you understand what scores mean and how to compare your participants' results to those of other groups. A standardized measure generally provides some security and credibility to your findings; it lends more rigor and is of particular value if you are interested in replicating your program or publishing your results.

Designing Your Own Tools

It is not the ideal, but many programs find that they need to develop their own measurement tools. Few programs have the resources needed to adequately test a new measure for validity, reliability, or other properties, but they are nevertheless able to create a tool that meets their needs. There are a few steps to be sure you include when you are developing a tool for use in your program.

Adapt an existing tool. You may have found something that seems almost appropriate. Perhaps it is a tool that measures the right concept and will fit into the context of your program, but it has not been translated into the language you need. Or, perhaps, it is almost right, but it is intended for an age group older than your target population. In these types of cases, your simplest course is probably to translate, adjust the language, or make other changes that address the barriers to using the tool in your program. When you change a tool, you lose advantages such as established validity and reliability, as well as the existence of comparison standards. However, you have the benefit that someone else has already done most of the work of figuring out what to ask. Be sure to talk to the author before adapting a tool. It is possible that someone else has already adapted the tool in a way that is better suited to your target population. You will also need to be sure that you have appropriate permission before you make any changes.

Review the literature. Even when you do not find an established tool in the literature in your field, it is helpful to know what is being written about in terms of both your intervention and the issue you are trying to address. You may also find information about what other programs like yours have tried to measure and whether anyone else has tried to look at the outcome or indicator of interest.

Talk to other programs. People who work in programs that are similar to yours, or who work with a similar target population, may have tools that will be useful to you. Their experience with data collection may provide some guidance, even when the programs are quite different. If you are a part of a network of programs, like the Healthy Tomorrows Partnership for Children Program, there is a good chance that someone else is dealing with the same issues that you are facing.[4] A Listserv in which you participate may also be a good place to get feedback or ideas on how to handle issues from your peers.

Talk to those with expertise or experience. National and local experts on the issue you are addressing, the population you are serving, and your community may have ideas about how to measure the outcomes your program seeks to achieve. Be sure to involve your program "experts"—your staff and recipients of your services—in identifying items to include in your tool.

Pilot test tools. ***Pilot testing*** refers to trying your tool with a few representatives of the group with whom the tool will be used. When the information is collected, you can check for issues such as difficulties in responding to particular items, incompleteness, and misunderstanding of directions or questions. You can see whether responses on your pilot tool make sense in relation to other information you may have or collect from your pilot test group.

---

[4] The AAP maintains a grants database that includes all Healthy Tomorrows grants as well as other grants administered by the AAP; this database can be accessed at http://www.aap.org/commpeds/grantsdatabase/grantsdb.cfm.

CASE STUDY

Sarah and the Prevention First Program staff realized that, for most of the information they needed to gather, there were no existing tools available to them. Staff would need to develop data collection tools for the Prevention First Program. Sarah expanded her chart to keep track of the tools they would need to use.

Prevention First Program Tools

Outcome	**Indicator**	**Source**	**Tool**
Immigrant families will understand the importance of preventive health care services.	Family rating of the importance of preventive health care after 6 months of program participation	Participating parents—possible survey	Family intake, participant survey
Participating immigrant families will schedule and complete an increased number of well-child visits.	Number of well-child visits among children from participating families in the first and second years of the program	Program records—family case records, visit logs	Contact sheet
Immunization rates will increase among children in the target population.	Change in immunization rate 2 years after program is implemented	Community-level immunization data from the county	(Collected by county)
The number of workdays or school days missed due to illness will decrease.	Participant report of missed days after 1 and after 2 years of program participation	Participating parents—possible survey, school records	Participant survey

It's your turn again! Have you been able to identify existing tools to measure one or more of your indicators? What tools will you need to develop for your program?

WORKSPACE

Outcome	**Indicator**	**Source**	**Tool**

Remember that you do not want to overwhelm your staff or your program participants with data gathering. If you identify multiple tools, you may need to make choices about which tools will work best with your target population and within the context of your program. You will also need to remain flexible if it becomes apparent that you are not getting the information you need from the tools you have chosen.

JARGON ALERT

!

Quantitative data

Quantitative data are data in a numerical format.

SIDEBAR

Using Qualitative Data in Evaluation

Qualitative data, simply defined, involve the collection of data in the form of words rather than numbers. The collection of qualitative data may be an appropriate part of your evaluation strategy. Some types of information cannot easily be obtained through other means, for example, answers to questions such as "Why?" and "How?" Often, qualitative information is used to supplement or help in the interpretation of ***quantitative data.*** Although a number of data collection methods can be used to gather qualitative data, those most commonly used in evaluation are in-depth interviews and focus/discussion groups. This type of data collection can allow you to:

- Develop insight into feelings, attitudes, opinions, and motivations.
- Preserve the chronological sequence of events or developments.
- Understand and describe the local context of your program.
- Study selected issues in depth and detail.
- Obtain the broadest possible response to a question or issue.
- Gather rich information about a small number of cases.
- Put a human face on the program.
- Get information about the language and terminology used by the target population and others in your community.
- Better describe the target population of the program, including challenges and strengths as well as needs.

Planning Data Collection

Like so many other things in life, effective information gathering requires good planning. Data collection works well when it is integrated with service delivery. All program staff must know who is collecting what information in what time frame. If data collection is at all complicated, or separate from service delivery, you may want to set up a tickler, or reminder, system to help staff keep track of what is due when. Your data collection plan must also be clear about *where* data will be collected and stored, particularly if there is some need for privacy or access to other resources; in addition, you need to know *how much time* data collection is likely to take.

SIDEBAR

Are You Going to Need Help?
You may decide that you need outside help for all or a part of your evaluation; in some cases this may be a requirement of your grant. Some things you might look for in an evaluator include:

- Specific evaluation training
- Applied research experience
- Experience in a human service setting
- Professional perspective and methodological orientation that match program needs
- Interpersonal style that fits into the program
- Self-interest (ie, the ability to put yours first!)

If you decide to hire outside help, be sure that you agree up front about who will do which tasks, the cost, and the timeline. Should you be sufficiently fortunate to have a university nearby with appropriately trained professionals, you may be able to find limited or low-cost help from a new researcher or well-supervised graduate student who is eager to find data to work with and possibly publish. In these cases, the same standards should apply, particularly in terms of meeting the information needs of the program first.

CASE STUDY

Sarah realized that she was becoming a little overwhelmed with all of the information she and her staff were talking about needing to collect. She sat down and created an overview of all of the tools they were planning to use to ensure that everyone was clear on the purpose of each. The staff reviewed the list together and selected the tools that were most important for their program implementation and evaluation. They decided to include the focus group they planned to conduct because it involved information gathering by the program. Sarah found her chart very helpful both in working with her staff and in talking to others in her organization. The overview was also potentially useful in setting up schedules and checklists for managing the collection of information by project staff.

Prevention First Program Data Collection Overview

Tool	Information	Timing	Completed By	Purpose
Process				
Focus group with immigrant parents	Barriers to use of prevention services	Once in first 6 months	Focus group facilitator	• Service planning
Family intake	Participant characteristics, health care history, status, resources, identified needs	Entry into services	Staff, with family	• Description of participant population • Examination of variations in program effects • Service planning
Contact sheet	Date, type of contact, location, result	Each individual family contact	Staff	• Description of service delivery and immediate results • Documentation of outputs (health care coverage)
Group participation log	Date, topic, attendees	Each session	Staff	• Description of service delivery
Outcome				
Participant survey	Participant status update, program experience	Every 6 months during program participation	Parent	• Documentation of outcomes (importance of prevention, missed days)

WORKSPACE

If you will be using a number of tools for various purposes, you may find it helpful to create a chart similar to Sarah's using the form below.

Data Collection Overview

Tool	Information	Timing	Completed By	Purpose

Assuming the schedule for collecting information is clear, you will want to keep in mind a few other guidelines that will make your life easier when you are setting up your system.

Keep your data collection:

Simple. The sophistication of your evaluation and, therefore of your data collection, should be appropriate for the scale of your program.

Focused. Do not collect any information you will not use, no matter how interesting it seems.

Ethical. Protect the privacy and dignity of your program participants and other respondents. In fact, you may find that you have very specific requirements with which you must comply if your organization has an Institutional Review Board (see Appendix B for more information on IRB).

Consistent. Information should be collected in the same manner for each person at each time point.

Collecting data can be a substantial effort for a program, especially if staff are already stretched thin. If resources are limited, consider collecting information from a ***random sample*** of your target population or other group that you are trying to represent in your data. Use the resources you have to do as thorough a job as possible, and you will have very good information on a representative sample rather than sketchy information on a self-selected group.

JARGON ALERT

!

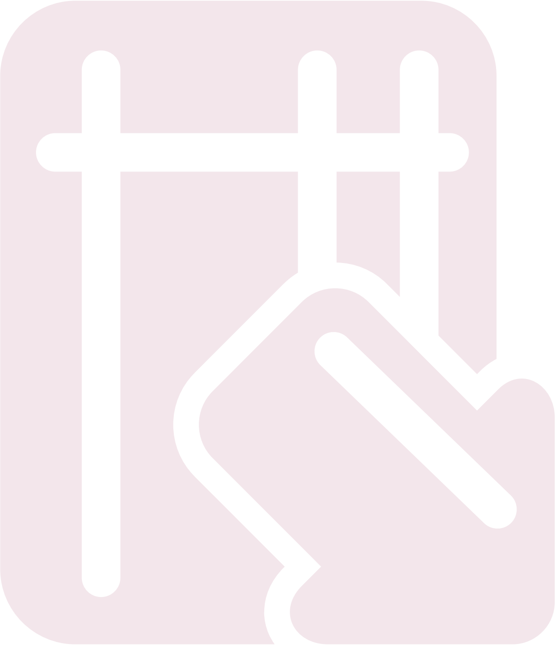
Random sample

A ***random sample*** is a group of subjects selected from a population of interest such that each member of the population has an equal chance of being included in the study group. Some familiar methods of simple random sampling include choosing names from a hat or flipping a coin for each case to determine inclusion.

Follow-up and Tracking

Sometimes, programs believe that they need to follow people over time to determine whether changes have occurred. Following people who are not receiving services can be very challenging and costly. Keeping track of people and collecting information from them incurs costs in terms of both dollars and staff time. Moreover, community-based programs often target individuals who face particular difficulties in terms of employment, housing, transportation, and other basic necessities. They may not have reliable telephone service or a long-term address. These vulnerable populations may be especially challenging to follow over time. For these reasons, it is probably best not to attempt what amounts to a follow-up study on a small program budget. One helpful option might be to collect exit information from program participants as they leave the program; this will allow you to gain the participants' perspective on how the program is working. This also may be a good time to use qualitative methods to explore the reasons people are not coming in or are hard to reach.

If you decide that you must follow people over time, tie information gathering as closely as possible to service delivery. If the program has any ongoing contact with people, those contacts are the best opportunity to find out how things are going. During the intake process, ask participants to provide contact information for someone "who will always know where you are." Try to maintain current contact information by periodically mailing out newsletters or other program materials. You can also maintain some contact by sending birthday cards or other appropriate messages to individuals even (or especially) when they are not receiving services from the program.



Managing Your Data

Data Handling and Storage

By this point, you should have a plan in place for collecting good information in a consistent manner. Sad to say, even excellent data collection is not sufficient. Very good data stacked in the corner of your office are not of much value to anyone. You need to keep the information organized and accessible for use in answering questions about your program and reporting to funders and other stakeholders.

Be clear about whose responsibility it is to manage data. If no one is assigned, chances are it won't get done. In most community-based programs, staff are focused on serving the needs of the target population, not on performing the evaluation activities. This is usually true even when staff are fully engaged and supportive of the evaluation.

Set up your system for storing information before you begin to collect it. If your information is relatively simple and straightforward, you can probably manage it all in a spreadsheet. This can be either a paper system or a computer software program such as Excel. The clear advantage of the latter is ease and flexibility when you compile your information. In addition, simple formats can usually be imported into data analytic software by a data analyst, should that become necessary or desirable. If your information is more complicated, involving multiple forms and time points, you will need to consider a ***relational database,*** such as Access, to keep it all organized.

Protect confidentiality. Individual case records should always be protected, starting with storage in locked files in the program office. To the extent possible, avoid entering identifying information into the same document record that includes sensitive information about an individual or family. Use passwords to restrict access to participants' private information. There may also be specific HIPAA or IRB requirements with which you must comply. (See Appendix B for specific information about HIPAA and IRB.)

JARGON ALERT

!

Rectangular databases

Rectangular databases require that each record include all variables (even if blank); an Excel spreadsheet is an example of this type of database.

Relational database

A ***relational database*** is one, such as Access, in which data are organized as a set of tables with defined relationships to one another. Information can be accessed or reorganized in many different ways. In contrast to rectangular databases, a table may include only the subset of cases for which the items it contains are relevant. For example, information about women may be held in one table and linked to a second table that includes child information only for those who are mothers.

Begin data entry immediately. Whatever system you are using, keep the information up to date. Enter data into the computer program or onto the paper log that is being used to store it. If it seems like too much to handle now, think about how overwhelming it will be if you need to produce a report from that stack in the corner!

CASE STUDY

Sarah planned to hire a data analyst when the time came, as neither she nor her staff members felt comfortable working with the data. However, Sarah knew from past experience that she should not wait to begin entering the information the program was collecting into a computer database. She had some familiarity with Excel and was able to set up a spreadsheet to organize the data. She labeled the columns for each of the variables being collected and used a numbering system for the individual families to protect their privacy. Following is a small sample of her database.

Family	***N* Kids**	**Country**	**English**	**Health Coverage**
1	3	USA	yes	yes
2	1	Mexico	yes	yes
3	5	Mexico	limited	no
4	2	Honduras	no	yes
5	4	Senegal	limited	yes
6	4	USA	yes	no
7	2	Mexico	no	yes
8	1	Serbia	no	no
9	3	Honduras	limited	no
10	4	Peru	yes	yes

Analyzing Your Data

Quality of Data

The usefulness of your information will depend, in large part, on its quality. Quality can be undermined by poor evaluation planning or implementation. Sometimes, data quality is affected by events beyond your control, such as a change in your community or a new state law. Some important issues in data quality are discussed here.

Representativeness. How well does your information represent the target group or community of interest? For example, if you followed up by telephone with program participants, you would miss anyone who does not have a telephone. The information you gathered would represent the experience and status of only those who had telephones, leaving out an entire group of people whose stories may be very different. This and similar gaps make it difficult to claim that the data you have in hand can be used to describe and assess your program.

Completeness. Do you have all of the information you intended to collect for everyone? Information may be missing for several reasons. Perhaps it was difficult to complete a particular tool at the correct time point, or maybe specific bits of information were not available for some cases. Incomplete information makes it difficult to describe your program and its effects.

Comprehensiveness. Did you collect information on all of the factors you want to include in your analysis and reports? Although simple data collection is best, it is difficult, if not impossible, to go back if you oversimplified. Your analysis and reports will be constrained by the information in your system.

Cleanliness. Are your data relatively free from errors and inconsistencies? You probably didn't know that your grandmother was talking about your data when she said, "Cleanliness is next to godliness!" If you work with a data analyst, he or she is likely to begin with data cleaning. This entails running some logical checks on the data, looking for things like people listed as being 203 years old, or someone who is coded as both having dropped out of high school and working as a public school teacher. Sometimes you will be able to help the analyst correct these kinds of errors, and other times people may have to be left out of a particular analysis because the errors cannot be corrected and would lead to confusion.

Data Analysis

Your data analysis may be very simple, involving only counting program participants who reached a milestone or calculating an average score on a posttest. Some programs with more resources may conduct analyses that require statistical expertise. Regardless, the rule is the same: your data analysis flows directly from your program evaluation questions.

SIDEBAR

Descriptive and Inferential Statistics

The field of statistics can be divided into two very broad categories: ***descriptive statistics*** and ***inferential statistics.*** Descriptive statistical methods are those generally used to organize and summarize numerical data collected from a group of interest, such as the proportion of the group that is female or their average score on a screening test. These methods meet most of the needs of community-based programs. Inferential statistical analysis seeks to draw a conclusion about a population based on a sample from that population, as, for example, health researchers do when they publish obesity rates among Americans based on a survey of a representative sample. In this type of analysis, we often test whether an observed difference or change is likely to have occurred by chance using a test of statistical significance appropriate for the data being analyzed, such as the comparison of two groups on a test score. Note, however, that such tests of ***statistical significance*** are not always necessary or appropriate to answer the questions you have about your program.

Keep it simple (again). Most program-specific questions are relatively simple: Are we reaching our target population? Are we achieving the outcome of interest? Do those who receive more service do better? For the most part, these questions do not require very complex analytic procedures. Don't get distracted by what may be interesting; focus on what you need to know and what you need to show your board, funders, and other stakeholders.

JARGON ALERT

!

Statistical significance

Statistical significance means that a finding is unlikely to have occurred by chance.

Computer-aided analysis, which offers so many options, may be especially distracting because data can be examined in so many different ways. Particularly if you have access to an outside evaluator or data analyst, who, after all, has interests and curiosity of his or her own, you may find yourself drawn to exploration of relationships in the data that are interesting but should be secondary to your program questions. This is not inherently a bad thing, but it will be your responsibility to be clear about your priorities.

When you approach your data analysis, you will be more focused if you are able to specify the questions you want to answer about your program; this is especially true if you are working with an outside evaluator or data analyst. Your evaluation plan has been based on your logic model. Your analytic questions will bring together the pieces of your model and your data collection to guide an analysis that uses the correct procedures and information in order to best answer those questions.[5]

CASE STUDY

Sarah realized that she needed to be specific about her questions regarding the Prevention First Program. She was going to be able to work with a data analyst on loan for a few days from another program, and she wanted to make the best possible use of that time. She also wanted to help the analyst find the information from her various data collection tools that was available to address her questions.

Analytic Question	Source of Information for Analysis (Tool)
Process Evaluation	
Has the Prevention First Program reached its target population in the immigrant community?	Family intake
What services do participants receive? Are they served primarily in traditional or nontraditional settings?	Contact sheet Group participation log
Outcome Evaluation	
How well have we achieved our outcomes?	Participant survey County immunization data
How do family characteristics affect program participation and outcomes?	Family intake Contact sheet Group participation log Participant survey
Can we identify a link between the level of program participation and level of outcome?	Contact sheet Group participation log Participant survey

---

[5] In another kind of research project, you would have begun by defining research questions to guide you in all of your decisions regarding evaluation planning and implementation. Your logic model has served this function in your evaluation project.

WORKSPACE

Your turn again. Can you identify the important analytic questions that will guide your data analysis? If you will be working with an analyst, consult with him or her when trying to specify your analytic questions. A skilled analyst should be helpful in refining the questions, and it will be useful to him or her to go through your thought process with you.

Analytic Question	Source of Information for Analysis (Tool)

A few simple procedures. Most of the analysis you will need to do will involve a few simple statistical procedures. All of these can be readily done by hand; however, you can also try them more ways, faster, and with fewer errors if you have access to a computer with Excel (a spreadsheet), SPSS (a statistical analysis package), or similar software.

If statistical analysis is completely beyond your comfort zone, this may be a good point at which to recruit some outside help. Many programs find that they benefit substantially from just a little bit of time from an experienced data analyst. An analyst can help you make sure that your data are analyzed correctly to answer your evaluation questions, and he or she can also help you to interpret the results.

Descriptive Analysis

Descriptive statistics in the form of percentages, averages, and graphs are generally used to organize and summarize the information you have collected. At a minimum, you will be expected to be able to provide information that describes your program, including who is receiving services and what services they are receiving. A good description of the population served and what services are used is often a strong starting point for talking to stakeholders and potential funders about your program. You will also likely want to be able to characterize the status of program participants when they complete services. You may find it useful to identify the characteristics of subgroups of your service population, such as those who have the highest success rate, or those who leave the program without completing services.

The most basic analytic procedure used in descriptive analysis is a ***frequency distribution.*** This is a simple tally of responses by category. In a computer-aided analysis, you will likely start by generating frequencies for all of the ***variables*** you have collected. This will form the basis of the answer for many of your program questions.

CASE STUDY

The data analysis conducted to answer Sarah's identified questions about the Prevention First Program used multiple analytic methods. From frequency distributions, she was able to see that recent immigrants comprised a large proportion of her participants and that most had limited or no English language skills.

Characteristics of Prevention First Program Participants: Country of Origin and English Language Skills

Country of Origin	**Percentage**
United States	22%
Mexico	37%
Honduras	22%
Senegal	9%
Other	10%
TOTAL	100%
English Language Skills of Parents	**Percentage**
Fluent in English	31%
Limited English	31%
No English	38%
TOTAL	100%

Comparing Groups

If you have the capacity to do more than simple descriptive analysis, you will find yourself making comparisons among the groups within your data. Different factors may affect the success of your program, and it is helpful to examine these in order to better deliver services to the people in your target population. Some of these factors are individual characteristics, such as age, gender, and ethnicity, and others are group characteristics, such as poverty and immigration status. One factor that may be very important to examine when you analyze your outcome data is the status of participants when they entered your program. Programs sometimes discover that participants are much needier than had been anticipated or that their immigration status is widely variant. Understanding where people are starting can help you tailor and target your services. It can also help you be realistic with funders and stakeholders in your community when setting goals and objectives for the future.

A simple procedure for comparing groups is ***cross-tabulation,*** sometimes called ***contingency table*** analysis. Like frequencies, a cross-tabulation is also a tally, but the tally is divided between groups such as girls and boys or those who received a well-child visit and those who did not. A cross-tabulation is most useful for comparing groups on variables with just a few categories, such as gender or ethnicity. The test most often used to see whether a difference in distribution among the categories between groups is statistically significant is a ***chi-square.*** Other reasons you might want separate tallies for groups include substantial differences in the number in each group, knowledge or suspicion that the groups had some difference at the beginning, and categories that may have different meanings for each group.

If you have a quantitative measurement or a score on a variable of interest, you can compare groups by ***comparison of means.*** A ***t-test*** is commonly used to test whether the difference between the means is statistically significant or unlikely to occur by chance.

JARGON ALERT

Categorical data

Categorical data puts subjects into a limited number of categories that are not necessarily ordered. Examples include gender, ethnicity, and political party affiliation. This type of variable is also called **nominal.**

CASE STUDY

Simple frequency distributions allowed Sarah to see that her program had a substantial rate of success in enrolling families in health care insurance programs (one of her program's outputs) and that many program participants completed a well-child visit (an outcome). Sarah's staff had suggested to her that it was particularly challenging to enroll families from Senegal in health care insurance. Sarah looked at her data and created a cross-tabulation to see whether this appeared to be a real issue and found that participants from two countries, Senegal and Honduras, seemed to be more challenging to enroll. Based on this information, Sarah can explore whether there are unique barriers these two groups face in enrolling in health care insurance programs and adjust the program as needed to better serve these families.

Prevention First Program Participating Families: Enrollment in Health Care Insurance by Country of Origin

Family Enrolled in Health Care Coverage by Prevention First Program Staff	Yes	No
United States	75%	25%
Mexico	55%	45%
Honduras	35%	65%
Senegal	20%	80%
Other	65%	35%

As the data analysis progressed, Sarah found that she was able to answer most of the questions she had identified. She and her staff were then prepared to work together and with other stakeholders to consider how best to use the information to better serve their target population and their community.

Change over Time

Making change happen is the reason your program exists. If your outcome of interest is something you assessed at the beginning of the program, you will be able to use this information in your analysis. You can look not only at where program participants ended up but at how much they changed from the beginning to the end of your data collection. This is particularly valuable if you have a target population with serious disadvantages to begin with, such that their "improved" status may not look particularly good when compared with the general population. For example, a weight loss program may be able to demonstrate significant improvement among its participants even if the participants are still obese after program participation.

There are some special analytic issues in looking at change in a group over time. The most important requirements are to have information at each time point of interest for all participants and to be able to match each participant across time points. The analysis then compares each participant to himself or herself. The simple statistical tests are the ***McNemar test*** for change (for cross-tabulations of ***categorical data*** between the two time points) and the ***paired t-test*** (to compare means on numerical variables at the two time points).[6]

Comparison Information

Although a formal control or comparison group may not be realistic on your program budget, you may have other comparison data that can be used in your analysis. Similar statistical tests can be used to test whether your participants are significantly different from or similar to the general population as represented by community, state, or national data or from an absolute standard. Less formal documentation of events and possible issues with ***bias*** and ***maturation*** can also be very useful and will help you interpret and explain findings when you report them. Characteristics of participants who either complete or drop out of your program may be factors to consider when comparing data.

JARGON ALERT

!

Bias

Bias is a flaw in evaluation design, sampling, or data collection and analysis that causes systematic error, possibly leading to incorrect conclusions.

Maturation

Maturation is a term used for the changes that occur in the program intervention group as a result of the passage of time and not because of the intervention itself.

Summary of Quantitative Statistical Procedures*

	Usual Procedure	Test of Significance	Change over Time
Numerical data			
1 group	Mean	NA	Paired *t*-test
2 groups†	Comparison of means	*t*-test	
Categorical (nominal) data			
1 group	Frequency distribution	NA	McNemar test
2 groups†	Cross-tabulation	Chi-square	

Shaded cells are not addressed in this guide.
*See the glossary for definitions of individual statistical terms.
†Procedures are similar for more than two groups.

---

[6] More complicated procedures, such as trend analysis and time series analysis, are available, but they are beyond the capacity and the need of most community-based programs.

Analyzing Qualitative Data

In some respects, analyzing qualitative data is more challenging than analyzing quantitative data because the process is more open-ended. However, as with quantitative analysis, your program evaluation questions drive the analysis. Some key considerations are that the analysis will ideally be conducted by multiple people who will check and balance one another and that the analysis should be verifiable by another party. Typically, the goal of qualitative analysis is to identify and analyze the themes that are present in all of the information that has been gathered. Usually, there is some form of validation involving staff and program participants. Frequently, programs that collect qualitative information do not go through a formal process of analysis. Instead, they use the information to provide stories and quotes to put a human face on the program and enrich their reporting.

SIDEBAR

Approaching Qualitative Data Analysis

In *Qualitative Data Analysis*, Miles and Huberman[7] outline three components of qualitative data analysis, all of which occur concurrently and inform the other components.

Data reduction involves reducing the words in interviews, field notes, and other sources to manageable "chunks" primarily through focusing and simplifying the data. Activities might include summarizing, coding, identifying themes, and associating data chunks within or across observations or subjects.

Data display is the way in which the analyst assembles and organizes the data for drawing conclusions. Narrative text is the most obvious form of display, but other displays, such as graphs, matrices, and charts, may be more useful.

Conclusion drawing/verification is the process of extracting meaning from the data. As the analyst notes patterns, sequences, and associations, he or she begins to draw tentative conclusions. The analyst works with colleagues and other data sources, if they exist, to test the validity of the conclusions, defined by Miles and Huberman as plausibility, sturdiness, and confirmability.

These steps are not so different from those used in quantitative data analysis. However, quantitative methods and the protocols for using them are well defined and are usually less cyclical than qualitative approaches to data. The authors point out that "humans are not very powerful as processors of large amounts of information." In other words, humans have a tendency to be selective, to focus on what seems simple or easy to understand, and to emphasize particularly vivid or distinctive information. Developing a process that is as objective as possible and documenting that process so that the analysis can be reproduced are critical to ensuring that qualitative analysis does not simply reflect the preexisting perspective of the analyst.

---

[7] Miles M, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Newbury Park, CA: Sage, 1994.

Using Your Findings

Making Decisions with Your Data

Before you use your data to make decisions about program adjustments, replication, and resource allocation, you need to decide whether you believe that your data can be trusted. First, you know how the process of evaluation planning and implementation has gone; are you satisfied with the integrity of the data? Second, given all that you know about your program, your community, and your target population, do the findings make sense to you? Can you explain the findings reasonably? A few other ways to explore whether to trust the findings are discussed here.

Look for consistency. Do the findings seem to relate to one another in expected ways? Two indicators that measure the same thing should change in the same direction. For example, if one measure of mental health shows improvement, a similar measure should also improve.

Look for trends. Does the progression over time make sense? For example, children should get larger and age should increase.

Ask questions. Particularly if you have the assistance of an evaluation consultant or data analyst, be sure that all of your questions are answered such that you understand and can explain the results.

Check with others. Review the findings with a representative group of everyone involved in the project, particularly including program staff and people from the target population. Ask one another whether the findings make sense and reflect together on what they mean.

SIDEBAR

Remember the Evaluation Feedback Loop

Now is the time to think about whether you need more or different information. Are any of your concerns with your data and findings reflective of problems with the evaluation plan or implementation? Is it possible that people are misunderstanding or misinterpreting questions or instructions? Do you need to modify your evaluation plan or measurement strategy?

Keep in mind that when you make a change in the information you collect or how you collect it, you lose the ability to compare information across time and make the best use of the information you have already gathered. If you are not getting good-quality information that you can use to answer your evaluation questions, you must consider making changes. However, the cost of less essential “tinkering” and refinements may not be worth the possible gains.

CASE STUDY

As they reviewed their evaluation findings, Sarah and the Prevention First Program staff realized that they needed to make a few adjustments to their evaluation plan. One source of information that they had planned to use, county data on community-level immunization rates, turned out to be unavailable in a form they could use. They also discovered that they needed to make more of an effort to record their contacts with individual families, as they believed the reported numbers underrepresented the work they were doing. They considered changing some of the questions on the participant survey to refine question wording and response options. They decided against these changes because the survey was generating information that was useful, and they did not want to lose the ability to compare the program and families across time.

SIDEBAR

Common Uses of Evaluation Information among Community-Based Programs[8]

- Improve services.
- Advocate for service population.
- Obtain funding.
- Support replication.
- Market services or organization.
- Promote policy change.

---

[8] Source: Healthy Tomorrows Partnership for Children Program: Highlights and Lessons Learned From the National Evaluation, 2006

Presenting Your Findings

Now that you have your data, think about with whom and how you want to share the information. At a minimum, you will want to get information from your evaluation back to your community and your target population. However, other groups may be interested in hearing about your program, including current funders, potential funders, community partners, and professional colleagues.

When you present your findings, you will need to provide a context for your data collection. Report how the information was collected, including any limitations or issues with data quality. This gives your audience the best opportunity to draw conclusions. It also helps keep audience members from being distracted by questions about the methods behind the findings.

Your presentation needs to be appropriate for its purpose and the target audience. A community forum is very different from a presentation to a funder or potential funders.

Simple messages work best. Know before you start what major point or points you want to make, and focus on those.

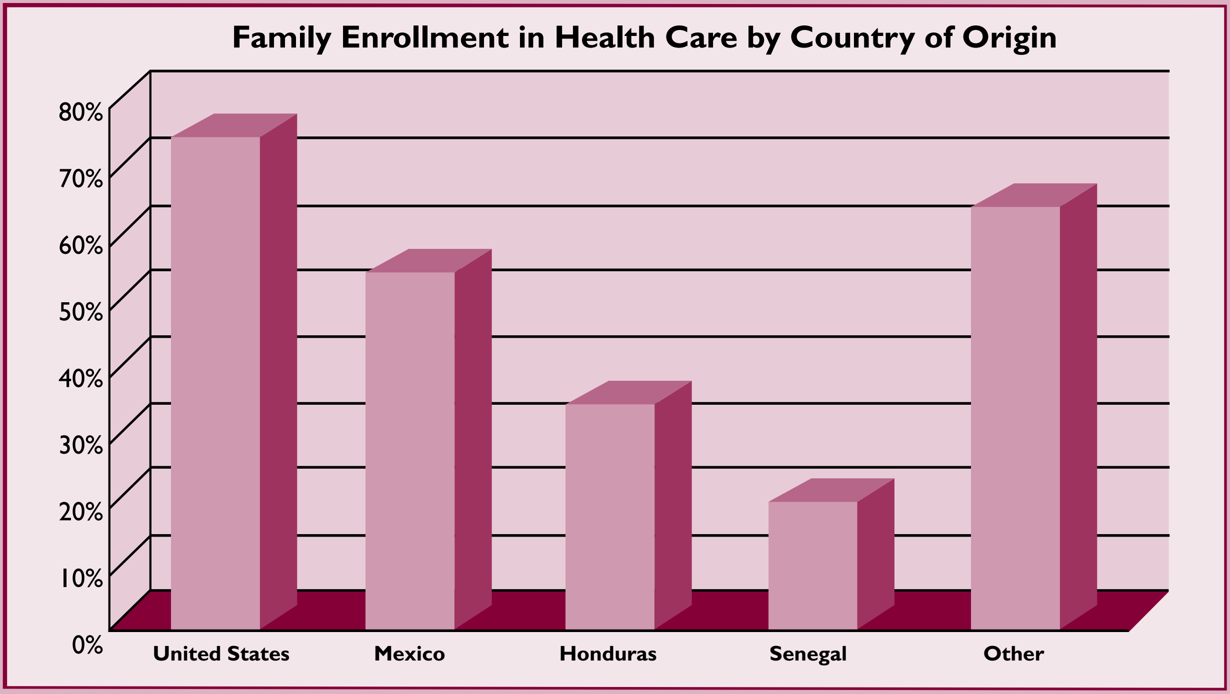
Match detail to the audience. Some groups will be genuinely interested in the details of your implementation process, but many will not. Think about the needs of your audience and, of course, what you need from them in advance.

Make presentations as visual as possible. Present text as brief bullet points. Use charts, graphs, tables, and pictures as much as possible. These presentation tools both make the presentation more interesting and help your audience absorb the information you want them to remember.

CASE STUDY

The data analyst had put most of the findings into tables for review by Sarah and the Prevention First Program staff. However, when it came time to present the findings to other audiences, Sarah felt that the tables were rather dull, and she was concerned that not all of the stakeholders would have an easy time reading the columns of numbers. Sarah created simple charts and graphs of the findings she wanted to highlight. The following sample chart visually represents the cross-tabulation Sarah created to examine health care enrollment by country of origin for the families served by the Prevention First Program. She remembered the focus group they had conducted at the beginning of the program, and she was able to enliven the data by incorporating a few quotations that illustrated some of her results. Sarah was able to make the presentation still more interesting by including some photographs of families the staff had taken during a community picnic in which the program had participated.

Prevention First Program Participating Families



Conclusion

We hope that this guide has helped you to:

- Identify ways to measure progress on your goals and objectives.
- Select tools and strategies for collecting information you need to evaluate your program.
- Think about how to analyze and present your information in a meaningful and useful way.

As in Part I of this guide, each of these topics could be a book in itself; in fact, books have been written about most of them! A few of these resources are listed in Appendix A of this guide. We have attempted to highlight the critical issues to consider when you are making decisions about collecting and analyzing information about your program. We have also tried to suggest ways in which you can use your data to serve the needs of your target population and your community. We hope that, in working through this guide, you have developed an evaluation plan that will answer the important questions about your program as well as some concrete tools to help you implement your plan.

HEALTHY
TOMORROWS
Partnership for Children

Glossary

Activities: Day-to-day ways in which people and material resources are used to achieve your goals (may also be called services, tasks, or strategies).

Baseline: Information about the target population or community collected before the intervention begins. Baseline information on indicators and outcomes is useful in examining whether there has been a change in the target population.

Bias: A flaw in evaluation design, sampling, or data collection and analysis that causes systematic error, possibly leading to incorrect conclusions.

Categorical data (also known as nominal data): A variable, such as gender or ethnicity, that has a limited number of categories. The order of the categories is not necessarily meaningful.

Chi-square: A test of statistical significance that is commonly used to compare observed and expected frequencies in a cross-tabulation. The larger the chi-square statistic, the more likely that the distribution is not due to chance, that is, the more likely that the variables in the table are associated.

Comparison information: Information drawn from existing sources or groups against which program information can be compared. Comparison information is useful for assessing how program participants are different from other populations, including how they differ on outcomes of interest.

Comparison of means: If there is a quantitative measurement or a score on a variable of interest, the means (averages) on the variable of the two groups can be compared. A test of significance, such as a *t*-test, tests whether the observed difference is likely to have occurred by chance.

Contingency table (also known as cross-tabulation): A table of counts, most commonly two-dimensional, showing the frequencies of two variables in relation to one another, displayed in rows and columns, respectively. It is most useful for comparing groups on variables with just a few categories, such as gender by political affiliation. The test most often used to see whether a difference in distribution among the categories is statistically significant is a chi-square.

Convenience sample: A comparison group drawn from another group in the community that is not receiving the intervention of interest but on whom data are or can be collected to compare with data from a group of program participants.

Cross-tabulation (also known as contingency table): A table of counts, most commonly two-dimensional, showing the frequencies of two variables in relation to one another, displayed in rows and columns, respectively. It is most useful for comparing groups on variables with just a few categories, such as gender by political affiliation. The test most often used to see whether a difference in distribution among the categories is statistically significant is a chi-square.

Descriptive statistics: Organization and summary of the information you have collected in the form of percentages, averages, and graphs.

Focus group: A professionally facilitated, focused discussion among a group of people to help understand a topic of interest. Usually, individual focus group participants are selected based on specific characteristics relevant to the topic.

Frequency distribution: A tally or count of the number of times each level or score on a variable occurs in a sample or population. A frequency distribution is easily presented as a bar graph.

Goal: A high-level, broad statement that articulates what a program would like to accomplish.

Impact: Change in a population, situation, or health issue targeted by your program.

Incidence: The number of cases of disease having their onset during a prescribed period. It is often expressed as a rate (eg, the incidence of measles per 1,000 children 5 to 15 years of age during a specified year). Incidence is a measure of morbidity or other events that occur within a specified period.

Indicator: A measurable intermediate step or other approximation of an outcome. An indicator is used when the outcome itself is difficult to measure directly and/or difficult to measure during the time frame of the project.

Inferential statistics: Allows an investigator to make an inference about a population based on a sample from that population by using a test of statistical significance to determine whether an observed difference or change is likely to have occurred by chance.

Input: The resources (human, financial, and other) available to your program to provide the services that will allow you to reach your desired goal.

Logic model: A visual representation of your program that illustrates the relationships among the resources available to you, what you plan to do with them, and your intended results.

Maturation: Changes that occur in the program intervention group as a result of the passage of time, not because of the intervention itself.

McNemar test: A special case of the chi-square test in which the cross-tabulation is of the same individuals on a single variable at two points in time. The McNemar tests for change between the two time points, rather than association between them.

Nominal data (also known as categorical data): A variable, such as gender or ethnicity, that has a limited number of categories. The order of the categories is not necessarily meaningful.

Objective: A measurable step toward achieving your goal.

Outcome: Measurable, intended results (short or long term) of your activities, strategies, and/or processes. May also be called impact, result, effect, or change resulting from your project.

Outcome evaluation: A plan to measure what difference your project is making for the target population.

Output: Basic information on participation or completion resulting from activities or services your project provides; used to measure or track the implementation process.

Paired *t*-test: A *t*-test statistic that compares scores at two points in time for the same group of individuals or other matched pairs.

Pilot testing: A trial run of your tool with a representative group of users. Piloting helps ensure that your questions and protocols are appropriate for your target population.

Prevalence: The number of cases of a disease, infected persons, or persons with some other attribute present during a particular time. It is often expressed as a rate (eg, the prevalence of diabetes per 1,000 persons during a year).

Process evaluation: A plan to measure whether your project is being implemented as you intended, including who is participating and what services are being delivered.

Program evaluation: A process for determining whether a project or program is achieving its goals and objectives.

Qualitative data: Data in the form of narrative or words providing detail and description rather than numbers.

Quantitative data: Data in a numerical format.

Random sample: A group of subjects selected from a population of interest such that each member of the population has an equal chance of being included in the study group. Some familiar methods of simple random sampling include choosing names from a hat or flipping a coin for each case to determine inclusion.

Randomized clinical trial: A research study utilizing a pool of subjects from a population of interest who are randomly assigned to treatment and control (no treatment) conditions.

Rectangular database: A database structure that requires that each record include all variables (even if blank); Excel is an example.

Relational database: A database in which data are organized as a set of tables with defined relationships to one another; Access is an example. In contrast to a rectangular database, not all variables need to be included for every case, and not every case needs to appear in every table.

Reliability: An assessment of the internal consistency of a measure. A reliable measure includes items that are related to one another and that give a consistent estimate of the construct being measured.

Selection bias: A specific type of bias introduced when the individuals who are selected for the intervention are in some way different from other people, especially from any comparison group. Any differences found may have been preexisting or related to factors other than the intervention. A special case of selection bias is caused by self-selection, when some people choose to join a group or seek services and others do not.

Standardization (of a measure): A measure or tool has been tested in one or more populations. Results can be compared across groups using percentages or other standard scores.

Statistical significance: A value is sufficiently large (or small) that it is unlikely to have occurred by chance.

Tool: In evaluation and research, a specific mechanism used to collect information, also sometimes known as an instrument or a measure.

***t*-test:** A test statistic most commonly used to assess whether the difference between two means (averages) is statistically significant, that is, unlikely to have occurred by chance.

Validity: A representation that a tool really measures the thing it is intended to measure.

Variable: A characteristic that can assume any of a set of values that is measured or assessed during data collection.

Appendix A: Evaluation Resources

Please note: Listing of resources does not imply an endorsement by the American Academy of Pediatrics (AAP). The AAP is not responsible for the content of the resources mentioned in this publication. Phone numbers and Web site addresses are as current as possible but may change at any time.

Finding Tools

1. **Published compendia of measures for specific topics or disciplines.** The following three resources are mentioned in the text of this guide; there are many others for diverse fields of study.

   Corcoran K, Fischer J. *Measures for Clinical Practice: A Sourcebook*. 3rd ed. Volume 1: *Couples, Families and Children*. New York, NY: The Free Press, 2000. The measures included are intended to assess change resulting from therapeutic interventions. Volume 2 deals with adults.

   Dahlber LL, Toal SB, Behrens CB. *Measuring Violence Related Attitudes, Beliefs and Behaviors Among Youths*. Atlanta, GA: The Centers for Disease Control and Prevention, 1998. This is a compendium of assessment tools from youth violence prevention programs funded by the Centers for Disease Control and Prevention in 1992–1993.

   McDowell I, Newell C. *Measuring Health: A Guide to Rating Scales and Questionnaires*. 2nd ed. Oxford: Oxford University Press, 1996. Focused explicitly on health indicators, this reference includes scales assessing general health status, physical disability, social health, psychological dimensions, pain, and quality of life.

2. **A relatively simple guide to the process of developing a measure.** Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. 2nd ed. Oxford: Oxford University Press, 1995. For those who wish to know more, this book is a good guide to understanding the process of instrument development and a generally practical resource for developing one's own scales.

③ **Web sites devoted to a specific issue or area of study. Never underestimate the power of your search engine (or your local health sciences librarian)! Here are a few examples:**

Agency for Healthcare Research and Quality (AHRQ) (www.ahrq.gov)
In particular, try these two potentially useful areas within this large site:

> ***Consumer Assessment of Healthcare Providers and Systems (CAHPS) (www.cahps.ahrq.gov/default.asp):*** look under "Survey Products" on the left for descriptions and downloadable versions of tools
>
> ***Child Health Care Quality Toolbox (www.ahrq.gov/chtoolbx)***

The AHRQ Web site offers information for consumers and practitioners, research reports and funding opportunities, and an extensive array of tools and resources. You could spend a lot of time following links through this Web site, although the menus and indices are generally clear and easy to follow.

American Diabetes Association (ADA) (www.diabetes.org)

The ADA is primarily an advocacy organization, but it provides links to relevant research (click on "Diabetes Research") on its Web site.

The Community Tool Box, University of Kansas (http://ctb.ku.edu)

This toolbox includes measures of community change among its tools to promote community health and development.

The Data Resource Center for Child & Adolescent Health (www.childhealthdata.org)

This Web site supports the national child health surveys conducted by the National Center for Health Statistics. It includes access to the data generated by the surveys, including information on state and regional levels, as well as the specific questions asked on the surveys. To get to the survey questions, click on one of the surveys and then find "survey information" in the menu on the left.

The Federal Interagency Forum on Child and Family Statistics (www.childstats.gov)

The forum describes itself as "a working group of federal agencies that collect, analyze, and report data on issues related to children and families." The ChildStats Web site facilitates access to statistics and reports from these agencies (as well as a few private partners). You will have to do a little work to find the actual tools, but start by clicking on "Data Sources" at the top, and then "Data Tools" on the next page that appears. You can then find links to agencies that collect information related to your topic of interest.

Medical Outcomes Trust (www.outcomes-trust.org/instruments.htm)

This is a listing of measures of general and condition-specific health functioning, with brief descriptions and information for obtaining the actual instruments.

Youth Obesity Learning Collaborative of the Association of Community Health Improvement (www.communityhlth.org/communityhlth/projects/youthobesity/obesityhome.html)

The collaborative provides links on its site to an extensive array of resources. Click on “Resources” on the left to see them. Access to the Community Assessment Toolkit (also in the menu on the left) requires membership in ACHI (or the American Hospital Association), but the cost of membership is relatively modest.

4 **Questionnaire Development Web-Based Teleconference (www.aap.org/commpeds/resources/teleconf_questionnaire.htm)**

The slides and the audio from this AAP Division of Community-based Initiatives teleconference provide a quick review of the basic steps in developing a questionnaire.

Data Analysis

1 **The University of Wisconsin-Extension: Using Excel for analyzing survey questionnaires (http://learningstore.uwex.edu/Using-Excel-for-Analyzing-Survey-Questionnaires-P1030C0.aspx)**

This is a step-by-step guide for conducting some very simple analyses using the Excel software you may already have on your computer.

2 **Online calculators to help you with some parts of your analysis**

Statistics Online Computational Resource (SOCR) (www.socr.ucla.edu)

This Web site provides extensive information about statistical analysis, including online calculators to perform statistical tests on your own data. A number of different departments at UCLA, including Statistics, collaborate on the Web site.

Interactive Statistical Pages project (http://statpages.org)

This Web site describes itself as “a comprehensive statistical analysis package that’s freely accessible over the Internet.” The site, which allows users to access information and links provided by volunteers from around the globe, has information to help you choose analytic procedures and interpret findings; it also includes statistical calculators for numerous tests.

3 A few relatively simple books about statistics

Gornick L, Wollcott S. *The Cartoon Guide to Statistics*. New York, NY: HarperCollins, 2005. The principles of statistics are presented graphically and with a sense of humor. If you want a general introduction and would prefer to pretend that you are not reading a statistics textbook, this may be a good choice for you!

Huff D. *How to Lie with Statistics*. New York, NY: WW Norton and Company, 1993. This slim gem will help you understand statistical reasoning and will put statistics in proper perspective. Originally published in the 1950s, some of the examples are a bit out of date, but mostly charmingly so. Note: There are no equations in this book!

Vogt WP. *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences.* Newbury Park, CA: Sage Publications, 1993. This is a reference to have handy if you find you must talk to a statistician. This book is very helpful for quick definitions of statistical terms and procedures.

Qualitative Data

1 The University of Wisconsin-Extension: Qualitative data analysis (http://learningstore.uwex.edu/Analyzing-Qualitative-Data-P1023C237.aspx)

The extension service specializes in very simple approaches to getting your questions answered.

2 A few comprehensible books on qualitative methods and analysis

Krueger R, King J. *Involving Community Members in Focus Groups*. Newbury Park, CA: Sage Publications, 1998. This volume focuses on working with non-researchers. It is part of the 6-volume Focus Group Kit edited by Richard Krueger and David Morgan.

Miles M, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed. Newbury Park, CA: Sage Publications, 1994. This book is very extensive but also relatively user-friendly.

Morse J, Richards L. *Read Me First for a Users Guide to Qualitative Methods*. 2nd ed. Newbury Park, CA: Sage Publications, 2007. This textbook provides a basic introduction to qualitative research.

Topics Included in the Evaluation Resources Appendix to *Evaluating Your Community-Based Program, Part I: Designing Your Evaluation*

(www.aap.org/EvalResources)

How to Evaluate
Logic Models
Community-Based Health Projects and Interventions
Public Health and Community Assessment Data
Evaluation Training

For additional evaluation resources and for updated Web site links, please visit the Community Pediatrics Evaluation Resources and Tools Web page at www.aap.org/commpeds/resources/evaluation.html.

Appendix B: HIPAA/IRB

Health Insurance Portability and Accountability Act (HIPAA) and Institutional Review Board (IRB)

Comply with standards for protection of client confidentiality. The Health Insurance Portability and Accountability Act (HIPAA) and the accompanying privacy regulations set national standards for protecting personal health information. The act allows health care providers to use patient data for program evaluation and quality improvement activities, which are considered health care operations. It also sets out the specific conditions under which data can be disclosed for research purposes and creates a "safe harbor" when data are de-identified by removing 18 specific identifiers: (1) names; (2) all geographic subdivisions smaller than a state, except for the initial three digits of the ZIP code if the geographic unit formed by combining all ZIP codes with the same three initial digits contains more than 20,000 people; (3) all elements of dates except year, and all ages over 89 or elements indicative of such age; (4) telephone numbers; (5) fax numbers; (6) e-mail addresses; (7) social security numbers; (8) medical record numbers; (9) health plan beneficiary numbers; (10) account numbers; (11) certificate or license numbers; (12) vehicle identifiers and license plate numbers; (13) device identifiers and serial numbers; (14) URLs; (15) IP addresses; (16) biometric identifiers; (17) full-face photographs and any comparable images; (18) any other unique, identifying characteristic or code, except as permitted for reidentification in the Privacy Rule. A common approach used within research and program evaluation is to make sure that any other personal identifiers beyond the HIPAA "safe harbor" are not linked in any way to the health information that is distributed outside the health care organization.

In addition, states often have their own laws governing privacy and confidentiality. All health care organizations that are covered by HIPAA are required to have a privacy officer. You should consult with your health care organization's privacy officer on your program evaluation plan. For specific guidelines on HIPAA and its use related to program evaluation and research, you can contact the US Department of Health and Human Services Office of Civil Rights.[9]

---

[9] The Web site address is www.hhs.gov/ocr/hipaa/privacy.html.

Comply with your organization's institutional review board guidelines, if applicable. Universities, research institutions, and some major health services organizations have an Institutional Review Board (IRB) in place to review all research and evaluation connected to that institution to ensure it complies with guidelines and standards. The federal government provides standards that are interpreted and applied by the local IRB in a manner consistent with institutional standards. In general, if you are reviewed by an IRB, the members will look for:

- *Risk/benefit analysis:* What are the risks to participants, and is the benefit generated from the evaluation greater than this risk?
- *Selection of subjects:* Are appropriate groups of participants targeted to answer the question proposed by the evaluation?
- *Informed consent:* Are participants adequately informed about their rights and the risks of participation as part of the ongoing consent process?
- *Privacy and confidentiality:* Are there adequate safeguards to protect participants' privacy in all data collection, analysis, and presentation?
- *Additional safeguards:* Are potential subjects protected from coercion or undue influence to participate, and are procedures in place to address potential harm to participants should it occur?

The IRB will also monitor the evaluation project as it proceeds, with particular attention to the collection, storage, and use of individual-level data. Some types of evaluation and research are considered exempt under the federal regulations, but this determination needs to be made by the IRB rather than by those conducting the evaluation or research project.

There are also some standards that apply to evaluation that do not necessarily apply to other endeavors using the scientific method.

Contact Information

Healthy Tomorrows Partnership for Children Program
Division of Community-based Initiatives
Department of Community, Chapter and State Affairs
American Academy of Pediatrics
141 Northwest Point Blvd
Elk Grove Village, IL 60007
847/434-4279
docbi@aap.org
www.aap.org/commpeds/htpcp

HEALTHY
TOMORROWS
Partnership for Children

A Cooperative Agreement Program of the Federal Maternal and
Child Health Bureau and the American Academy of Pediatrics

American Academy of Pediatrics
DEDICATED TO THE HEALTH OF ALL CHILDREN™