

CONCERNS ABOUT MONITORING AND EVALUATING ARH PROGRAMS



CHAPTER AT A GLANCE

- ▶ Reviews challenges to and offers tips on measuring the effectiveness of youth programs
- ▶ Discusses how to be sure that your results are attributable to the program effort
- ▶ Previews ways this Guide can provide information and offer support

Fifteen Challenges in Monitoring and Evaluating Youth Programs

1. Some MIS are not set up to track the special characteristics of youth programs.


Some MIS are part of a larger program or service delivery intervention. For example, a family planning program that has a youth component may be set up to track the distribution of contraceptives; it may not be set up to track services that are more likely to be utilized by youth, such as counseling or distribution of information, education and communication (IEC) materials.

Adapting your MIS to monitor an ARH program may require only minor modification, such as adding the specification of age in program utilization reporting. However, for larger-scale programs that reach groups other than youth, adding even one new component to the system may be difficult to institutionalize.

2. Tracking services does not guarantee that you will know how many youth you are reaching.

All programs need to determine how they will count the youth they are reaching and how knowing the number of youth reached will improve performance. Many programs count services, such as the number of meetings held or the number of condoms distributed. However, if all you know is that you distributed 1,000 condoms, you will not know whether 100 youth received 10 condoms each or 500 youth received 2 condoms each. Your information tracking system should try to collect key characteristics of program participants to help assess whether the program is reaching the number and type of youth it was designed to reach.

Collecting information about target population characteristics will also help you understand how your program participants change over time. For example, in the beginning, your program may target older youth, but as word spreads about the services available, your program may find



15 Challenges in Monitoring and Evaluating Youth Programs

1. Some MIS are not set up to track the special characteristics of youth programs.
2. Tracking services does not guarantee that you will know how many youth you are reaching.
3. You may be unsure whether general standards or implementation strategies are applicable in the country you work in.
4. Little is known about whether standards for adult programs are appropriate for youth.
5. The elements of successful youth programs have not been well-documented or disseminated.
6. Programs may have trouble developing systems that understand and respond to the needs of youth.
7. Measuring the quality of a program requires understanding complex meanings and addressing sensitive issues.
8. Measuring a program's access and coverage can be complex.
9. Assessing individual reactions to a program can be difficult.
10. Measuring influences on behaviors that didn't occur is difficult.
11. Measuring behaviors at a variety of developmental levels can be problematic.
12. Showing the link between health outcomes and youth development can be complex.
13. Some changes may not be measurable for a long time, and others may be hard to measure at all.
14. Attributing changes in outcomes to a particular program's strategy and activities is difficult.
15. Some types of evaluation are costly and may require funds beyond a youth program's resources.

itself working with younger adolescents and need to adjust its approach accordingly.

3. You may be unsure whether general standards or implementation strategies are applicable in the country you work in.

Quality refers to the appropriateness of a specific set of professional activities in relation to the objectives they are intended to serve.¹ Standards of quality for the design of health education programs have been drawn from a variety of youth programs demonstrated to be effective in changing specific behaviors² and include factors such as:

- a minimum of 14 hours of instruction,
- small groups and an interactive environment, and
- models of and practice in communication, negotiation and other skills.

However, we do not know the extent to which these standards apply in a more diverse set of developing country settings.

The recommendations in this Guide, such as the Logic Model described in Chapter 2, are designed to help you implement your program strategy, based on

¹ Green and Lewis, 1986.

² Kirby et al., 1997.

assumptions about the social and behavioral factors that influence the health outcomes you hope to produce. The theories these recommendations draw on are well-developed and have been through a rigorous process to test how well their measurements capture the processes of change they propose. Yet most of these theories have not been tested in developing country settings and need to be adapted to the particular needs of youth in each locale. Since program activities drive the design of any evaluation effort, our lack of understanding about how these theories apply in different contexts can also affect our ability to undertake solid outcome and impact evaluations.

4. Little is known about whether standards for adult programs are appropriate for youth.

After years of developing contraceptive service delivery systems for adults, there are now more or less accepted standards of quality. For example, there is wide consensus that the delivery of quality clinical contraceptive services entails:

- technical competence of service providers,
- respectful treatment of clients,
- effective communication with clients,
- choice of methods,
- mechanisms to encourage continuity, and
- cultural appropriateness and acceptability of services.³

³ Bruce, 1990.

⁴ Birdthistle and Vince-Whitman, 1997; Israel and Nagano, 1997; Senderowitz, 1997a; and Senderowitz, 1997b. Note that these key elements reflect the experiences of programs that are concerned more with reproductive health outcomes than with youth development outcomes.

However, we still do not know how comprehensive these standards are for younger age groups. Some of these quality standards are listed in the Indicator Tables as examples of criteria to include in indicators of quality, especially at the design stage.

5. The elements of successful youth programs have not been well-documented or disseminated.

Youth program staff in developing countries often must rely on intuition and experience to design their programs when they don't have access to documented research. However, much is known about the standards that produce effective programs. For example, the FOCUS on Young Adults program has identified the following "key elements":⁴

- baseline assessment conducted to identify issues, needs and target audiences;
- existence of a clearly defined mission statement that contributes to the achievement of program goals; and
- local stakeholders involved in program planning.

6. Programs may have trouble developing systems that understand and respond to the needs of youth.

Program systems and their functioning will influence factors such as staff performance, service delivery and program utilization. Program systems must be set up to respond to the special needs of young people. For example, the staff recruitment and training system must ensure that staff hold the characteristics and skills to which youth respond well. A program system will help identify whether program materials are being updated often enough to respond to the changing language and trends of youth culture. A training system must ensure that the necessary components of youth programming are included in the curricula.

7. Measuring the quality of a program requires understanding complex meanings and addressing sensitive issues.

To determine program quality, you will probably have to elicit subjective interpretations, perspectives and meanings from young people and others in the community. These are each complex because they are based on:

- ▶ cultural beliefs and values,
- ▶ personal interactions within a community,
- ▶ interactions between the young people and the program’s staff, and
- ▶ opinions and views of people carrying out the program.

Programs that are concerned with youth empowerment, community mobilization, changing social norms and influencing youth culture will need to explore the

elicit and harder still to quantify. For example, you may be able to count the number of community members at a meeting, but have more difficulty assessing their substantive contribution to the meeting, increased concern as a result of the meeting or proposed strategy for social change.

Substantive changes in meanings and perceptions are extremely important for youth programs and should not be minimized. They play an important role in the quality of a youth program. To capture these nuances, we need to first employ qualitative approaches to data collection. Once we understand the relevant meanings, values and beliefs we can then collect data about changes in the number of participants who share those meanings, values and beliefs, i.e., a quantitative approach.

There are numerous obstacles to measuring the outcomes of youth development and reproductive health programs, which helps explain why we have such a limited body of evidence as to “what works.” First, many of the intended outcomes are regarded as personal and private. In some societies, talking about sexual behavior and personal relationships may be socially prohibited. Second, evaluators may face parental and community resistance to asking young people questions. Community leaders or other key stakeholders may believe that the young people in their communities do not engage in risky behaviors, and therefore there is no need to ask questions. They may also find it socially or politically dangerous to uncover the truth about young people’s sexual behavior, and make an attempt to block data collection. However, there are many examples of programs that asked sensitive questions and found young people who were eager to discuss issues of sexuality and reproductive health—viewing the discussions as an opportunity for

Your information tracking system should try to collect key characteristics of program participants to assess whether the program is reaching the number and type of youth it was designed to reach.

meanings of such issues as feelings of self-worth, the value of community connectedness and the interpretation of culture. These reflections may be difficult to

learning and for sharing their own concerns and needs.

8. Measuring a program's access and coverage can be complex.

Access to reproductive health programs concerns the extent to which youth can obtain appropriate reproductive health services at a level of effort and cost that is both acceptable to and within the means of a large majority of youth in a given population.⁵ We can define access in a variety of ways:

- ▶ **Geographic/Physical:** Convenient hours and location, wide range of necessary services
- ▶ **Economic:** Affordable fees
- ▶ **Psycho-social:** Perception of privacy; perception that both males and females, married and unmarried youth, are welcome; feeling of safety and confidentiality; perception that providers are interested in, informed about and responsive to youth needs
- ▶ **Administrative:** Specially trained staff with respect for young people, adequate time for interactions, youth involvement in design and continuing feedback, short waiting times

Coverage refers to the extent to which your program's services—such as educational or clinical services—are being used by your intended target population. Coverage can be measured by:

- ▶ determining the proportion of the target population you are reaching, or
- ▶ determining the characteristics of the population you are reaching.

Some aspects of accessibility and coverage can be measured by the absence or presence of something and may be relatively straightforward. For example, finding out whether your program has

convenient hours and affordable fees may be easily determined with a short survey of your target population. However, measuring more subjective issues that involve judgments—such as whether staff have respect for young people—can be more difficult because many youth may be reluctant to give their true opinions about program staff for fear of negative consequences, such as having services withheld.

Similarly, determining some characteristics of youth may be simple, such as asking participants about their age, sex and place of residence. However, if your program is reaching specific groups of youth, especially those who are marginalized, it may be more difficult to collect these data. For example, if your program is attempting to reach youth who have been sexually abused, the subject may be too sensitive for participants to respond easily to questions. You may have to ask questions repeatedly and to reassure participants that it is safe to talk.

9. Assessing individual reactions to a program can be difficult.

One measure of quality is how your program is received by stakeholders, staff and youth participants. Assessing how the program is received by these groups will contribute to your understanding of how to overcome social resistance to youth programs. It will also help you determine if your program is headed in the right direction and identify problems in time to correct them. However, eliciting and analyzing individual reactions to programs is difficult to do.

For example, you may want to engage youth and community members to think critically about their needs and to consider how the program could best reach them. Yet, some individuals may have trouble articulating their needs, or their opinions

⁵ Bertrand et al., 1994.

may defy what we know about the factors that influence health outcomes. Some community members think it is dangerous to give reproductive health information to youth, and they may want to censor the media in order to produce positive health outcomes among youth. Others may automatically express views that are in line with social norms and values, even if these views do not reflect the true needs of the community. Youth, in particular, may be reluctant to express negative feedback about the program to evaluators, who are often older and carry more authority.

Measuring the social and cultural context of youth development is difficult and may require time and resources that many programs do not have.

Similar tendencies may be found in the reactions of program staff and volunteers. Process evaluations encourage staff to reflect on their work, to see its strengths and weaknesses and to consider alternative strategies. Yet, while most people working with youth are deeply concerned and committed, some have a more ideological approach. They may assume that their strategies are working, even if there is little evidence to suggest that this is true. For example, some staff may insist that increasing access to contraceptive services is the best way to produce results, ignoring the fact that for youth who are abstinent, a more important service may be support in reflecting on and supporting a decision not to have sex. Others may think that their

commitment and hard work should pay off in results, and find it demoralizing to discuss how their efforts may be misguided. Staff will need a trusting environment and a supportive process to allow for the kind of reflection in which they can admit that program strategies might need modification.

10. Measuring influences on behaviors that didn't occur is difficult.

Many ARH programs are concerned with preventing unhealthy behaviors and influencing developmental pathways. They are often concerned with measuring events that did not occur because of the program intervention. For example, some programs may aim to delay the onset of sexual activity or prevent unwanted sex. Others may try to prevent early marriage, thus attempting to delay young women's first sexual experience and increase the age at first birth to a time when delivery will be safer. Obviously, measuring the absence of certain behaviors is complex. It requires estimating what level of behavior would have existed had there not been an intervention, then explaining why an intervention caused behaviors not to occur.

11. Measuring behaviors at a variety of developmental levels can be problematic.

Although youth programs are concerned with reaching young people throughout a developmental transition, we are not always sure what outcomes should be expected at specific ages. For example, we may be unsure of what the average age at first sex in our target population is. However, measuring outcomes on sexual behavior can be problematic. Some young people may not have heard about certain sexual behaviors and therefore have problems answering questions about them. This could bias results (e.g., when a girl who has held hands with a boy reports that she has engaged in "sexual activity"). Community

members, and sometimes program staff themselves, may believe it is not appropriate to introduce youth to new topics, such as sexual behavior or illegal behaviors, through a data collection effort.

12. Showing the link between health outcomes and youth development can be complex.

Many programs are increasingly concerned with linking health outcomes to youth development. For example, a program may want to demonstrate that increasing girls' education helps to delay first sex and thus has a positive health outcome. However, what aspects of youth development influence health outcomes may be difficult to predict. We cannot assume that developmental factors would have the same influence on health in different settings, as outcomes are embedded in specific and local contexts, each with their own social and cultural values. Measuring the social and cultural context of youth development is difficult and may require time and resources that many programs do not have.

13. Some changes may not be measurable for a long time, and others may be hard to measure at all.

It may be several years before you can observe changes in the health status of young people, as opposed to the relatively short amount of time it takes to observe such outcomes as changes in levels of knowledge. Moreover, some changes in outcomes may occur long after the program is over; for example, a program that promotes delay of first sex among youth ages 10 to 12 may not be able to observe its results for several years after participants take part in the program. It is therefore important to track trends in such behaviors. For many of the outcomes we are concerned with, we do not know how long it will take to bring about changes. Yet, many youth programs are expected to

demonstrate changes in longer-term outcomes in a very short period of time. Some programs define their objectives unrealistically and then falsely conclude that the program did not succeed, when, in fact, more time was required to demonstrate the changes.

Similarly, some program strategies, particularly those that deal with social change, are difficult to measure in numerical or quantifiable terms. For example, measuring complex social processes, such as community mobilization and empowerment, can be difficult because conceptually we are not exactly sure how to define these processes, nor articulate how they are occurring.

Community leaders or other key stakeholders may believe that young people in their communities do not engage in risky behaviors, so they feel there is no need to ask questions.

14. Attributing changes in outcomes to a particular program's strategy and activities is difficult.

How can you conclude that the changes you observe in your target population occurred as a result of your program activities? Measuring changes in outcomes alone is not enough to conclude that the changes occurred as a result of your

program. Other events, like shifting economic or social conditions, could have affected the outcomes you are measuring. There may also have been other program activities directed at your target audience, such as a mass media campaign, going on at the same time. Finally, your program could have attracted participants who were predisposed to the positive outcomes you were trying to encourage. The primary way to determine that an observed change in outcome indicators is attributable to your program is to use a strong study design (see Chapter 5). However, planning and implementing a strong study design requires a high level of resources and skills and may not be feasible for some programs.

15. Some types of evaluation may require funds beyond a youth program's resources.

Outcome and impact evaluations can be costly, especially when measuring numerous outcomes or those that are more difficult to assess. If programs cannot rely on existing data sources, they may need to collect quite a bit of new information about the youth populations they reach. Developing survey instruments, conducting correct sampling procedures and collecting data from individuals can all be expensive. Programs that do not have in-house evaluation expertise may also have the added cost of technical assistance or hiring external evaluators.



13 Tips for Addressing the Challenges of Monitoring and Evaluating Youth Programs

1. Monitor what your system is set up to deliver: programs for youth.
 2. Base your program activities, and thus your evaluation effort, on theory.
 3. Review what is known about the factors that influence health outcomes.
 4. Test and document the elements that contribute to your program's effectiveness.
 5. Engage in a genuinely participatory process.
 6. Ensure that your data collection effort addresses ethical concerns.
 7. Be creative in asking sensitive questions.
 8. Define your objectives realistically and provide enough time to measure changes.
 9. Use a combined qualitative-quantitative approach.
 10. Use monitoring and process evaluation data to support the outcome and impact evaluation.
 11. Learn by trial and error.
 12. Limit evaluation costs when possible.
 13. Build on the advantages of evaluating youth programs.
-

Thirteen Tips for Addressing the Challenges of Monitoring and Evaluating Youth Programs

Program staff and evaluators around the world are honing their skills and developing creative solutions to the tough challenges of monitoring and evaluation. Below are tips from practitioners in the field and suggestions on how to use this Guide to address measurement challenges.

1. Monitor what your system is set up to deliver: programs for youth.

Monitor the elements of your program's system that respond specifically to the needs of youth. In the Indicator Tables in Chapter 10, we provide some notes on how you can develop a new system or adapt an existing system to capture the needs of youth programs.

2. Base your program activities, and thus your evaluation effort, on theory.

Basing program strategies on theory helps articulate how programs are working and, if they are successful, aids in their replication and adaptation. The Logic Model introduced in Chapter 2 is an example of how a program can plan its activities based on theories of health behavior and social change. Increased understanding of how these and other theories apply in different contexts will strengthen our ability to undertake scientifically sound outcome and impact evaluations.

3. Review what is known about the factors that influence health outcomes.

To help you demonstrate the link between health outcomes and development needs, you should first review what is known about the influences that you assume will affect outcomes. As Chapter 2 suggests, the best way to do this is to review the existing research and literature about your target

population. However, if you are unable to access the published literature, or if it is not well-developed in your setting, you can review your staff's experience or talk to colleagues from other organizations. You may also find ways to assess these influences through creative data collection, such as asking questions about a particular behavior in a number of different ways, or modifying language and terminology on your survey instruments to reflect the most important issues in your setting.

Some programs define their objectives unrealistically, and then falsely conclude that the program did not succeed, when, in fact, more time is required to demonstrate the changes.

4. Test and document the elements that contribute to your program's effectiveness.

A number of elements contribute to a youth program's effectiveness. For example, the design elements proposed in the Indicator Tables reflect the current state of knowledge about the design features that are key to program success. We suggest that these elements be tested (i.e., used and evaluated to find out whether they are appropriate and effective) or modified, according to your specific setting and program priorities.

The systems you set up in order to implement a youth program are also key to program success and may be more complex than the systems used to implement other reproductive health service programs. Moreover, the criteria for assessing the quality of youth program system functioning have not been systematically tested. Possible criteria for assessing the quality of your system and its operation are suggested in the Indicator Tables, such as:

- recruitment of staff with appropriate skills,
- components of the training program,
- training program participants who have mastered skills,
- content of reproductive health curricula, and/or
- staff performance.

A participatory process also encourages the community and staff to utilize the information from process evaluations and have a sense of ownership of evaluation results.

Incorporating the elements identified by the international experience of youth programs, lessons from the field of family planning and your own intuition and experience is the best way to establish quality programs and services for young people.

5. Engage in a genuinely participatory process.

Evaluation that engages and involves stakeholders and staff is more likely to produce reactions that are critical and honest than those conducted exclusively by external experts. A participatory process also encourages the community and staff to utilize the information from process evaluations and have a sense of ownership of evaluation results. Giving youth and adults the opportunity to discuss and analyze their concerns, and to suggest and enact solutions, may also increase your program's effectiveness in reaching its objectives.⁶ Tips for engaging youth, community members and other stakeholders are provided throughout this Guide.

6. Ensure that your data collection effort addresses ethical concerns.

Professional standards of conduct as well as moral principles and values should be exercised in conducting research and evaluation studies. Ethical reviews are designed to consider and mediate the potential risks and negative consequences to participants as a result of their participation in a study or evaluation. Responding to ethical concerns will improve your relationship with the community and enhance your ability to collect quality data. The more ethical your data collection effort, the more honest and reliable the information you collect, which ensures that your M&E results are valid. Strategies for ethical data collection among young people are discussed in Chapter 7.

⁶The literature on participatory process evaluation is well-developed. One resource specific to the context of young adult reproductive health programs is Shah et al., 1999. *Listening to Young Voices: Facilitating Participatory Appraisals on Reproductive Health with Adolescents*. Washington, DC: CARE International in Zambia and FOCUS on Young Adults.

7. Be creative in asking sensitive questions.

Asking questions of a sensitive nature, while difficult, can be done successfully in many different settings. First, you may need to get support from a broad range of community organizations, to whom you will need to make clear why these questions must be asked. Second, you will need to obtain parental consent, particularly for youth who are legally minors. Third, you can employ “skip patterns” to avoid sexually explicit questions about contraceptive use or other sexual practices if youth have not had sex. Additional tips are provided in the discussion on data collection in Chapter 7.

8. Define your objectives realistically and provide enough time to measure changes.

Programs that define objectives unrealistically may lead to false conclusions. You should budget plenty of time before attempting to measure changes in outcomes, and ensure that your objectives clearly state the outcomes that you expect to produce.

9. Use a combined qualitative-quantitative approach.

Qualitative methods can be used to define social and cultural contexts and develop vocabularies for health education programs, each of which contributes to the formulation of instruments to be used during quantitative surveys. Quantitative methods ensure standardized data collection over time and enable definitive measurement of changes in outcomes that can be generalized to the larger population. They can also be used to show that changes are due to your program activities. Qualitative data can then be used to interpret the findings of quantitative surveys and may reveal program results not discovered through quantitative methods.

Qualitative methods can also be used to assess program goals that are difficult to measure quantitatively, such as empowerment and social change. For example, one qualitative approach asks staff and participants to describe the evolution of the program. Employing this method can help us understand what changes were brought about and why. The results of this approach can then be used to develop a quantitative approach to measure whether those changes are producing the intended outcome in the larger community.

Measuring the absence of certain behaviors is complex.

In Chapter 5, we suggest using a combination of qualitative and quantitative approaches to develop indicators and collect data, which will help you address some of these concerns in more detail.

10. Use monitoring and process evaluation data to support the outcome and impact evaluation.

Conducting outcome and impact evaluations requires resources and time, and even those that are well-designed may not show conclusive results. Using monitoring and process evaluation data can strengthen the results of your outcome and impact evaluations. For example, your MIS may collect information about exposure to the program’s services, such as contact with peer educators. If you are trying to demonstrate such outcomes as the increased use of condoms, you may want to measure whether youth received condoms or referrals from a peer educator. While this will not give you conclusive evidence about

your program's effect on the entire target population, it may help you demonstrate associations between positive outcomes and exposure to your program's activities.

While the challenges described in this chapter are many, the task of evaluating youth reproductive health programs can be very gratifying.

The information you collect through monitoring and process evaluation will also help you build the case that the changes were a result of your program, even if an impact evaluation is not feasible. For example, your MIS may show that certain activities were carried out more frequently than others. Your process evaluation may determine that young people liked certain messages better than others and became more involved in the program as a result. It may also document that community leaders' support for specific activities resulted in increased participation or the addition of new activities. Documenting factors such as service utilization, program participation and reaction to program strategies will strengthen the case that your program produced the desired outcomes.

11. Learn by trial and error.

For other measurement concerns, we need to learn by trial and error. For example, we are learning that we can ask questions

about sexual behavior, even in settings with very traditional values. Who asks the questions, how we ask the questions, or the place in which we ask the questions may need to be modified in each setting. We also need to be creative about generating valid self-reports of risk behaviors, as we often get the "socially desirable" response rather than an accurate account.

At this stage, many of the suggested measurements in this guide have not yet been tested. We will build on what we know as we collect more evidence and as programs like yours undertake more systematic approaches to monitoring and evaluation.

12. Limit evaluation costs when possible.

While outcome and impact evaluation can be costly, there are ways for programs to limit costs and still produce valid results. For example, an evaluation can examine only those outcomes most important to your program. Measuring outcomes that require less costly data collection methods or utilize already-existing data can also reduce costs. Training and utilizing staff to conduct some parts of the evaluation may be feasible for some programs. A sound sampling strategy can help you limit the amount of data collected while not compromising the validity of your evaluation results.

13. Build on the advantages of evaluating youth programs.

While the challenges described in this chapter are many, the task of evaluating youth reproductive health programs can be very gratifying. There may be other fairly simple ways to avoid challenges in measuring outcomes, including randomly assigning youth in school settings, either individually or by classrooms, and following them. You may find communities where the demand for the program

outpaces the supply. You may also choose to evaluate the impact of selected programs, rather than all programs, or to use a delayed treatment design, which is discussed in Chapter 5. If a few youth are in desperate need of the program, allow them to participate in the program but not the study.

There are many advantages to working with youth. They are interested in learning, and changes in this population can occur relatively quickly. Youth are in a period of great vulnerability, and improving outcomes for them is an investment in our future.



PHOTO: Cathryn Wilcox, JHU/CCP

